PEPTIDOMIMETICS TO MIMIC PROTEIN-PROTEIN INTERACTIONS

A Dissertation

by

ZEBIN XIA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2004

Major Subject: Interdisciplinary Engineering

PEPTIDOMIMETICS TO MIMIC PROTEIN-PROTEIN INTERACTIONS

A Dissertation

by

ZEBIN XIA

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

James C. Rock (Co-Chair of Committee)	
Kevin Burgess (Co-Chair of Committee)	John P. Wagner (Member)
Andrew K. Chan (Member)	Karen Butler-Purry (Head of Department)

May 2004

Major Subject: Interdisciplinary Engineering

ABSTRACT

Peptidomimetics to Mimic Protein-Protein Interactions.
(May 2004)
Zebin Xia, B.S., Hunan Normal University, P.R.China;
M.S., Nankai University, P.R.China
Co-Chairs of Advisory Committee: Dr. James C. Rock
Dr. Kevin Burgess

Quenched Molecular Dynamics (QMD) used to explore molecular conformations was developed to operate in Insight II platform for two simulation engines: CHARMm and Discover. Two scripts and procedures were written for molecular minimization, dynamics, minimization of each of several hundred conformers, and cut off. Experience with Insight II/Discover versus Quanta/CHARMm, and between Insight II/CHARMm versus Quanta/CHARMm has taught that the forcefield is the key factor in QMD studies.

Protein A has been used for the purification of commercial antibodies, but it is expensive. Seven peptidomimetics of protein A were designed based on the hot-spots located at the helix-loop-helix region of protein A, and synthesized via solid phase using the Fmoc approach. These peptidomimetics were characterized by MS and NMR. The conformations of four peptidomimetics were studied by NMR and CD in water/hexafluoroisopropanol (pH 4). The CD and NMR data show that addition of hexafluoroisopropanol stabilizes their □-helical conformations. The structures of these peptidomimetics in solution were generated with Quanta/CHARMm using NMR data as limits for the QMD technique.

Protein G has also been used to purify antibodies, but it is expensive too. A number of protein G mimics were designed as trivalent molecules. An efficient preparation of trivalent molecules having a useful primary amine arm has been

developed through solid phase synthesis. The cheap, commercially available poly(propylene imine) dendrimers were used as scaffolds which allow multimerization of functionalized compounds. A small library of trivalent compounds were synthesized using this approach. A portion of compounds in this library were tested by Amersham Biosciences. The seven amino acid modified DAB-Am-4 exhibits strong binding to the IgG/Fab, and is a potential ligand for IgG purification.

The interactions between neurotrophins (*ie* NGF and NT-3) and their receptors are typical drug targets. Fourteen second-generation peptidomimetics showing NGF-like or NT3-like activities in a preliminary bioassay, were resynthesized and tested again. Preliminary and retested data were compared. To access a direct binding assay, five fluorescently labeled peptidomimetics **41a-e** were synthesized for a fluorescence activated cell sorting (FACScan) assay. Six monomeric precursors **42** and **43** were prepared on large scales for the library of bivalent turn analogs.

DEDICATION

I would like to dedicate this dissertation to my wife, my parents and my parents-in-law.

ACKNOWLEDGMENTS

Fist, I would like to thank my advisors: Dr. James C. Rock, for his useful guidance in my graduate study and for his kindness, and Dr. Kevin Burgess, for providing me the opportunity and excellent environment conducive to independent research and hard work. I would like to thank Dr. Andrew K. Chan and Dr. John P. Wagner for serving on my graduate advisory committee.

I would like to thank Dr. Karl M. Koshlap for his help in my 2D NMR spectra, Dr. Lisa M. Pérez for her help in the QMD exploration on Insight II platform. I would like to thank Miguel Angel Castañeda and Jingyi Shen for proofreading of my dissertation. I would like to thank everybody in the Burgess group especially Sam, Reyes and Mookda, Pattarawarapan for their friendship and helpful discussions.

I would like to specially thank my wife, Wenhong Zhu. I could not complete my study without her love and support.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Page

Page

Page

LIST OF SCHEMES

Page

LIST OF TABLES

Page

LIST OF ABBREVIATIONS

Boc	tert-butoxycarbonyl-
Fmoc	9-fluorenylmethoxycarbonyl-
DCM	dichloromethane
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
Aib	□-aminoisobutyric acid
Acp	aminocaprionic acid
HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	*N*- hydroxybenzotriazole
DIEA	di-iso-propylethylamine
TFA	trifluoroacetic acid
TIS	tri-iso-propylsilane
DIC	*N*,*N'*-Diisopropylcarbodiimide
TFFH	tetramethyl fluoro formamidinium hexafluorophosphate
DAB-Am-n	poly(propylene imine)dendrimer
Tris	tris(hydroxymethyl)aminomethane
HFIP	hexafluoroisopropanol
IgG	immunoglobulin G
RP-HPLC	reverse phase high performance liquid chromatography
QMD	quenched molecular dynamics
SD	steepest decent
ABNR	adopted-basis Newton Raphson
CD	circular dichroism

STD	saturation transfer difference
NOESY	nuclear Overhauser enhancement spectroscopy
ROESY	rotating-frame Overhauser enhancement spectroscopy
DQF-COSY	double quantum filter correlation spectroscopy
TOCSY	clean total correlation spectroscopy

CHAPTER I
INTRODUCTION: FUNDAMENTALS OF PROTEIN-PROTEIN INTERACTIONS

1.1 Importance in Medicine and Biotechnology

Protein-protein interactions play a critical role in numerous biological processes in cells. They are important in the normal function of signal transduction,[1-3] immune response,[4,5] protein enzyme inhibitors[6,7] etc. The abnormal protein-protein interactions are key factors in the development of some pathological processes, for example, Alzheimer's disease,[8,9] anemia, rheumatoid arthritis, and respiratory distress syndrome and emphysema.[6,7] Since data from the human genome project were interpreted, a huge amount of information about protein-protein interactions has been accumulated. A lot of protein-protein interactions become pharmaceutical targets. This not only provides great opportunities, but also are big challenges for medicinal chemistry and biotechnology.[10] First, most of protein-protein interactions are quite specific; second, protein-protein interactions *in vivo* are sensitive to both the levels of proteins and how they are distributed;[11] third, if the protein-protein interfaces consist of noncontinuous binding epitopes it is difficult to rationally design binding-site mimics; fourth, the interfaces in protein-protein interactions are commonly greater than 1200 $Å^2$,[12,13] which vastly exceed the potential binding area of small molecular mimics, and are often planar lacking binding sites for small molecules.[14]

This dissertation follows the style of the *Journal of Organic Chemistry*.

Despite of the difficulties in the mimic of protein-protein interactions, a number of small molecular mimics, which can be used to modulate protein-protein interactions, have been identified, and new marketed products and recent research success,[15] which target the protein-protein interactions, have been reviewed.[10,16-18] The following strategy is often followed in the design of new bioactive compounds:[19] **a**. identify protein-protein interactions, **b**. propose hot-spots of contact area, **c**. design a library of small molecules with molecular modeling, **d**. prepare the compounds, and **f**. screen bioactivities of compounds.

1.2 "Hot-spots" at Protein-protein Interfaces

An excellent understanding of protein-protein interfaces is required for intelligent design of small molecules which are able to interfere with protein-protein interactions. Most interfaces consist of two relatively large protein surfaces with good shape complementarity to each other through steric, hydrophobic and electrostatic interactions, and hydrogen bonds.[12,20-22] The interface size differs from study to study. Janin[23] reported that it ranges from 670 to 4890 $Å^2$ per subunit in dimeric proteins. Later, a range of interface size from 368.1 to 4746.1 $Å^2$ was found by checking 32 non-homologous dimers.[6] A more recent examination of 75 protein-protein complexes exhibited that the interface areas range from 1140 to 4660 $Å^2$, and 70% of these complexes have the interface areas burying 1600 (±400) $Å^2$, which were called "standard-size" interfaces.[24] Generally, larger than 1200 $Å^2$ of total surface area, 600 A^2 per monomer, is required for the formation of "O-ring", which excludes bulk solvent and provides a low dielectric environment for "hot-spots" binding.[25,26] Only a small number of "hot-spot" residues on protein surfaces contribute significantly to the binding, and they clustered near the geometric center of the protein-protein interface.[26] A survey of amino acids on hot-spot regions showed that tryptophan, tyrosine and arginine are enriched in this area.[26] Although the identification of hot-spots remains an open question, the following numerous methods have proven useful for this task: X-ray crystallography,

NMR spectroscopy, mutagenesis screening,[4,26-28] site-directed mutagenesis,[29,30] and others.[31-34]

1.3 Examples of Peptide-like Molecules that Mimic Hot-spots

Impressive progress has been made to date in the discovery of small molecules modulating protein-protein interactions. The design of small molecule inhibitors for various enzymes has been quite successful, and the structural properties of enzyme interfaces are probably the main reason for this. Most enzymes have deep pockets within their interfacial surface area unlike many other protein heterodimers whose interfaces are large and flat. Moreover, these deep pockets are often the active sites of the enzymes. Small molecular inhibitors of HIV-1 protease dimerization are good examples of peptide-like molecules mimicking hot-spots. The interface of HIV-1 protease dimerization consists mainly of an interdigitating *C*- and *N*- terminal four-stranded antiparallel □-sheets.[35] Specifically, about half of it is only composed of the area near the *C*- and *N*- termini, and accounts for close to 75% of the total binding energy of homodimer.[36] The following residues located at the *C*- and *N*- termini were identified as hot-spots by calorimetric experiments:[36] Cys95, Thr96, Leu97, Asn98, Phe99, Pro1, Ile3 and Leu5. Based on these hot-spots Chmielewski's group[37] designed their first peptidomimetic **A** (Figure 1.1) as an inhibitor of HIV-1 protease with an IC_{50} of 350 nM (at 25 nM protease). They then applied alanine scanning, Zhang-Poorman analysis[38] and molecular modeling to peptidomimetic **A**. After the removal of less-important amino acids and the replacement of the isobutyl group of Leu with a cyclohexyl group in peptidomimetic **A**, a lead compound **B** (Figure 1.1) with smaller molecular weight was developed, whose IC_{50} is 680 nM (at 25 nM protease).[39]

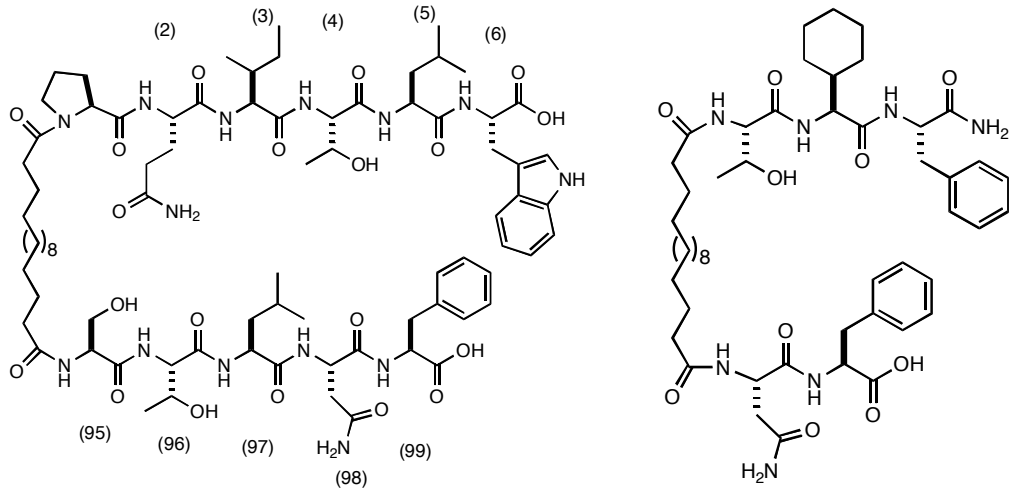
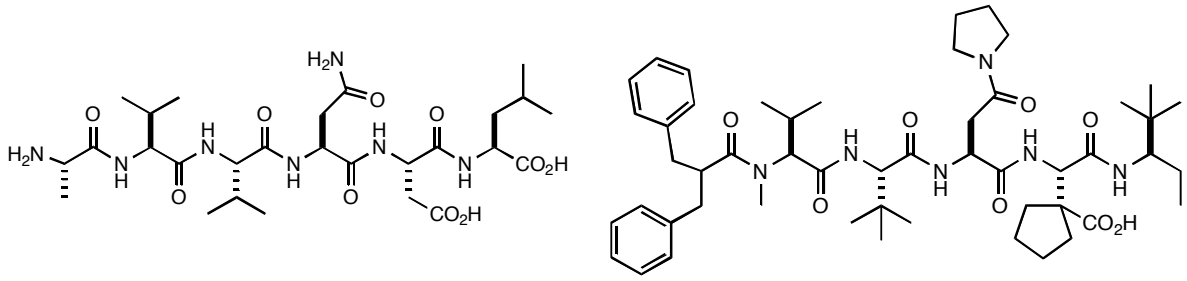


Figure 1.1. Structures of peptidomimetics **A** and **B**. The numbers correspond to the position of the residue in the complementary protease monomer.

Another excellent example is the mimicking of the hot-spots of the small subunit R2, which results in the active form of ribonucleotide reductase by associating the large subunit R1, and to inhibit herpes simplex virus ribonucleotide reductase dimerization.[40,41] Hot-spots clustered together on the C-terminal fragment (called hexapeptide **C**, IC_{50} = 58 μM) of HSV R2 (Figure 1.2).[42] Systematically varying side-chains of hexapeptide **C** led to the discovery of **D** (IC_{50} < 1 nM), which efficiently inhibited HSV replication in tissue culture.[43-45]



C **D**

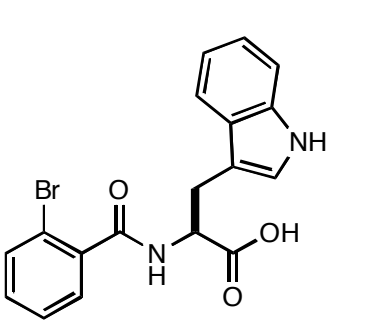
Figure 1.2. Structures of peptidomimetics **C** and **D**.

1.4 Examples of Non-peptide Small Molecules that Mimic Hot-spots

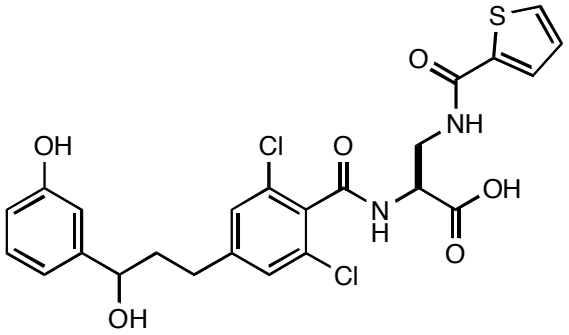
Mimics could be divided into two types: peptide and non-peptide small molecules. Beside the above peptide-like mimics there are many examples of non-peptide small molecules that mimic hot-spots modulating the protein-protein interactions. The following example successfully transformed the epitope of intercellular adhesion molecule-1 (ICAM-1) into a small molecule that is an effective antagonist of leukocyte functional antigen-1 (LFA-1). This epitope is composed of hot-spots in ICAM-1's first domain: E34, K39, M64, Y66, N68 and Q73.[46,47] The interactions between LFC-1 and ICAM-1 are pivotal to lymphocyte and immune system function.[48,49] Therefore, the development of new antagonists, which are able to be used to inhibit the binding of the LFA-1 and ICAM-1, block the lymphocyte function, and regulate the human T cell-mediated inflammation, is significant. ICAM-1, a LFA-1's native ligand, was employed as a lead and starting point through kistrin which contains RGD sequence and blocks LFA-1 – ICAM-1 interactions *in vitro*,[50] the RGDMP peptides,[51] H_2N-CGY$^{(m)}$DMPC-COOH, and compound **E** (Figure 1.3) (IC_{50} = 0.70 ± 0.21 μM for the inhibition of LFA-1/ICAM-1 binding), which was previously identified as an inhibitor of LFA-1,[52] during the identification of compound **F** (Figure 1.3) with IC_{50} = 0.0014 ± 0.00014 μM for the inhibition of LFA-1/ICAM-1 binding and IC_{50} = 0.003 ± 0.002 μM

for the inhibition of mixed lymphocyte reaction, using alanine scanning, molecular modeling and other approaches.[53] The molecular modeling shows that functional groups of the ICAM-1 epitope, such as carboxylic acid (E34), sulfide (M64), phenol (Y66) and carboxamide (N68, Q73), are embodied in compound **F** well.

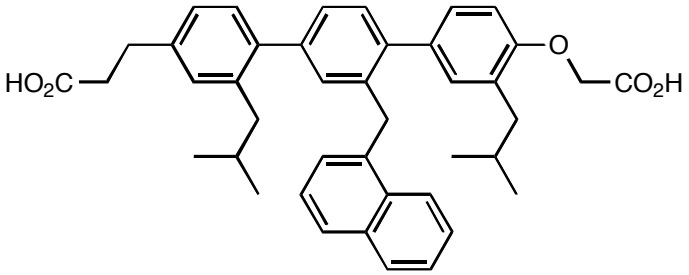
The development of a potent Bcl-x_L antagonist, which is a non-peptide small molecule, is another example of mimicking hot-spots. Bcl-x_L is a kind of anti-apoptotic protein while Bak is a type of pro-apoptotic protein.[54] The interactions of Bcl-x_L/Bak compete for supremacy over cell life and death decisions. In case of cancer Bcl-x_L is often overexpressed and protects transformed cells from cell death leading to uncontrolled cell growth in despite of the presence of apoptotic signals.[55] The crystal and solution structures of Bcl-x_L/Bak complex displayed that Bak binds the B_H1-B_H3 domains of Bcl-x_L via its helical fragment.[56,57] Amino acid residues Val74, Leu78, Ile81, Asp83 and Ile85 of the helical fragment of Bak were identified as hot-spots in the binding of Bcl-x_L.[58,59] Based on these hot-spots and the helix secondary structure a library of terphenyl molecules was designed. Of them, compound **G** (Figure 1.3) (K_d = 114 nM) with a helical structure was identified as a lead, and is a promising antagonist of Bcl-x_L.[60]



E



F



G

Figure 1.3. Structures of compounds **E**, **F** and **G**.

CHAPTER II
ADAPTATION OF INSIGHT II/DISCOVER TO CONFORMATION STUDIES OF PEPTIDOMIMETICS

2.1 Specific Aims

Computational molecular modeling has proven to be a very useful tool in modern drug discovery. The quenched molecular dynamics (QMD) is an excellent approach for molecular conformational exploration in the computational modeling. This project is to explore the QMD technique based on Insight II platform. Specifically, the goals of this work are to:

a. write two scripts and two procedures for molecular minimization, dynamics, minimization of each of 600 conformers, and cut off; one for Insight II/Discover while the other for Insight II/CHARMm;
a. link the two scripts to the main program of Insight II;
b. validate the scripts in Quanta and Insight II platforms using a specific molecule **H**, and compare the results;
c. perform QMD studies of the four stereoisomers of **P27**.

2.2 Background and Significance

Since Hendrickson[61] first stated that the energy of a molecule might be calculated on a computer in 1961, the field of computational chemistry has been rapidly extended to the study of peptides and other biological molecules.[62,63] These studies focus on the structures, energies, vibrational frequencies, and molecular dynamics simulations from small molecules to large molecules in pure liquid, or vacuum, or solution, or a crystal environment.[64-67] Recent advances in computer hardware, interactive graphics, and theoretical algorithms are opening up a new era in this field.

2.2.1 *Ab* Initio and Simulation Methods

Computational methods applied in chemistry and biology fields can be categorized into three classes: semiempirical, empirical and *ab* initio. Semiempirical method introduces significant simplifications which make it computationally much more feasible. Most integrals are neglected, and empirical parameters have to be introduced to compensate for such an approximation. Empirical methods differ greatly from *ab* initio and semiempirical methods since it is not based on the solution of the Schrödinger equation. Molecules are treated as the systems which are composed of atoms held together by bonds. The total energy of the system is expressed as the sum of bond stretching, bending, torsion, and attraction and repulsion between nonbonded atoms. *Ab* initio methods, also called nonempirical method, are based on the solution of time-independent Schrödinger equation, and is free of any empirical procedures and parametrizations. This method might be considered as a true theory. Despite that, it still includes some approximations, such as separation of nuclear and electronic motion (Born-Oppenheimer approximation), neglect of relativistic effects and concept of Molecular Orbitals (MO's). Since 1994[68] the application of ab initio method with inclusion of electron correlation has significantly changed the common view on interactions of molecules. It can be used to calculate not only electrostatic energy, dispersion energy, repulsion, polarization, but also electron correlation (e.g. Hartree-Fock method), allowing for reliable comparison of the strength of stacked and hydrogen bonded pairs of nucleic acid base, and characterization of the nature of the base-base interactions.[69-71] Hence, the ab initio study qualitatively improved our knowledge of the hydrogen bonding interactions which is important in life science, and revealed the intrinsic nonplanarity of amino groups of DNA.[72,73]

Molecular dynamics (MD) and Monte Carlo are very popular simulation methods in *ab* initio calculations. MD describes a movement process of a molecule in time, in which the motions of each atom including atom positions, masses and velocities are governed by Newton's equations. Monte Carlo method is based on a random

generation of an ensemble of conformations, or on statistical mechanics. Both are used for the samplings of conformational space.[74]

2.2.2 Force Fields

The force fields employ a combination of internal coordinates and special terms (bond distances, bond angles, torsions, etc.), to describe a part of the potential energy surface due to interactions between bonded atoms, and nonbond terms to describe the van der Waals and electrostatic interactions between atoms. The functional forms range from simple quadratic forms to Morse functions, Fourier expansions, Lennard-Jones potentials, and so on. The goal of a forcefield is to describe an entire class of molecules with reasonable accuracy. Some forcefields aim at high accuracy for a limited set of element types, thus enabling good prediction of many molecular properties. The forcefield contains necessary building blocks for the calculations of energy and force:

a. list of atom types
b. list of atomic charges (if not included in the atom-type information).
c. atom-typing rules.
d. functional forms for the components of the energy expression.
e. parameters for the function terms.
f. rules for generating parameters that have not been explicitly defined for some forcefields.
g. a defined way of assigning functional forms and parameters for some forcefields.

CFF and CHARMm are two important forcefields. CFF is capable of predicting many properties, and parameterized against a wide range of experimental observables for organic compounds containing H, C, N, O, S, P, halogen atoms, alkali metal cations, and several biochemically important divalent metal cations. CFF has been shown to reproduce experimental results more accurately than classical forcefields such as CVFF and AMBER.

CHARMm force field is mainly applied for biochemistry, and packaged in

a highly flexible molecular mechanics and dynamics engine originally developed in the laboratory of Dr. Martin Karplus at Harvard University[75]. It has been widely used and can be considered well tested and characterized. A variety of systems, from isolated small molecules to solvated complexes of large biological macromolecules, can be simulated using CHARMm.

Like CFF and CHARMm forcefields, AMBER and CVFF (Consistent Valence Force Field) are also applied for biochemistry. The AMBER force field was parameterized against a limited number of organic models. It has been mainly used in the study of proteins, DNA. CVFF is a classic forcefield having some anharmonic and cross term enhancements. As the traditional default forcefield in the Discover program, CVFF has been used extensively. Both AMBER and CVFF can be considered well characterized.

2.2.3 Methods Based on Molecular Dynamics

Molecular dynamics (MD) is one of the most important computational techniques. It can be used to generate a realistic picture of a structure's motions, perform conformational searching, do a time series analysis of structural and energetic properties, explore energy decay, and analyze solvent effects.[76,77] A molecular dynamics calculation generates a dynamics trajectory consisting of a set of frames of coordinates and velocities that represent the trajectory of the atoms over time. Using trajectory data, one can compute the average structure and analyze fluctuations of geometric parameters, thermodynamics properties, and time-dependent processes of the molecule.

In CHARMm and Discover, molecular dynamic simulation can be performed using a classical mechanics approach, in which Newton's equations of motion are integrated for all atoms in the system. With energy evaluations, a defined .psf or .car file, a set of coordinates and parameters are required to initiate a molecular dynamics calculation. The Verlet algorithm is one of the most often used approaches in MD.[78]

Several methods have been applied in molecular computation, such as QMD, simulated annealing, consensus dynamics. QMD has been used for many years in our

lab. Molecular simulation with this method is performed at high temperature. The structures generated at this temperature are periodically saved, and quenched via a form of minimization to produce a conformational ensemble of structures.[79] A typical QMD calculation used in our lab involves the following steps:

(1). preliminary preparation (in Discover)

A molecular structure in which all Cartesian coordinates are defined is required for a dynamics simulation. As the internal coordinate values of the molecule are determined, total energy of the molecule is computed by evaluating the individual terms of the energy equation.

(2). minimization

Energy minimization should be performed on structures prior to dynamics in order to relax the conformation and remove steric overlap that produces bad contacts. A minimized ideal geometry may then be used as a starting point for dynamics.

(3). heating

A minimized structure represents the molecule at absolute zero or a temperature near absolute zero. Heating starts at this temperature, and is accomplished by assigning greater random velocities to each atom at predetermined time intervals according to a Gaussian distribution appropriate for that low temperature. The temperature is then gradually increased.

(4). equilibration

At equilibration the average temperature and structure remain stable, and various statistical properties of the system become independent of time. Equilibration was achieved by allowing the system to evolve spontaneously for a period of time, and by integrating the equations of motion. This is facilitated by periodically reassigning velocities appropriate to the desired temperature.

(5). simulation

CHARMm and Discover take the equilibrated structure as their starting point. The trajectory traces the motions of the molecule through a period of at least 10 picoseconds. Conformers are then sampled as many as you want from the trajectory.

(6). quenching

This step is the logical opposite of heating; it takes the molecule from the equilibrated temperature to absolute zero. In fact, quenching is a form of minimization by utilizing molecular dynamics to slowly remove all kinetic energy from the system.

(7). cut off

It is used to remove the structures with higher energy, and keep the lower energy conformations for the next step.

(8). clustering families

The remaining structures via cutoff are clustered into a few families Based on the RMSD values, which are indications of the relative energies and conformational properties.

2.2.4 CHARMm and Discover

Both CHARMm and Discover are simulation engines including forcefields for minimization, dynamics, and other molecular mechanics simulations. CHARMm [75], which is the abbreviation of Chemistry at Harvard: Macromolecular Mechanics, is a large program designed for manipulation of structures and calculation of molecular properties. It facilitates comparison of structures and evaluation of conformational energies. It may also be used to minimize energies, perform a normal mode or molecular dynamics simulation, and analyze the structural, equilibrium, and dynamic properties determined in calculations. CHARMm can be used to deal with isolated molecules, molecules in solutions, and molecules in crystalline solids. It is also available to carry out a wide range of analysis such as static structure and energy analysis, structure and energy comparisons, correlation functions and statistical properties of molecular dynamic trajectories, and interfaces to computer graphics programs.

Discover is another large program written in C and FORTRAN language. It has been applied in life science more widely than CHARMm, and can be used to perform energy minimization, template forcing, torsion forcing and dynamic trajectories, and to

calculate properties such as interaction energies, derivatives, mean square displacements, and vibrational frequencies. Discover provides tools for performing simulations under various conditions including constant temperature, constant energy, constant pressure, constant stress, periodic boundaries, and fixed and restrained atoms.
Both CHARMm and Discover may run in a standalone mode. In this case, the following three files: an input file (run_name.inp, where run_name is your name for the calculation) containing special commands to control the calculation, and two files describing the model (run_name .crd or .mdf and run_name .psf or .car) are required.

The differences in minimization methods[80] between CHARMm and Discovery were shown in Table 2.1.

Table 2.1. Minimization Methods in CHARMm and Discover

Method	Variant	Simulation Engine	
		CHARMm	Discover
Steepest Descents		yes	yes
Conjugate-Gradient	Polak-Ribiere	no	yes
	Fletcher-Reevs	yes	yes
	Powell	yes	no
Newton-Raphson	Full, iterative	yes	yes
	BFGS(quasi)	no	yes
	DFP(quasi)	no	yes
	Truncated	no	yes
	ABNR	yes	no

2.2.5 Significance

Since β-turns and helices are pivotal to many protein-protein interactions our group has been interested in the types of β-turn mimics in which the residues of amino acids and synthetic templates are held together to mimic the conformations of the loop regions of the neurotrophins, and of α-helical mimics in which short peptidomimetics

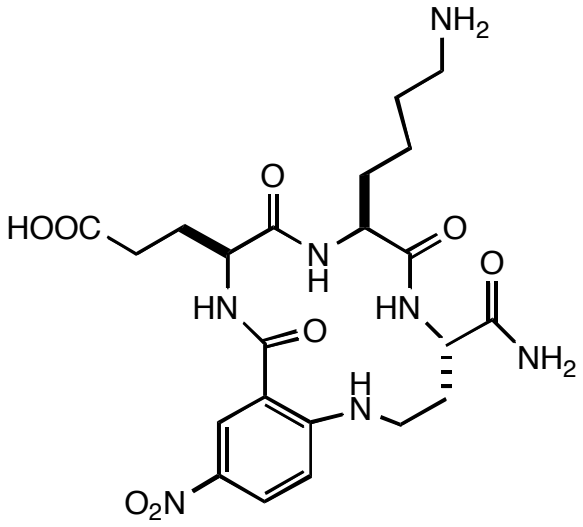
consisting of natural and unnatural amino acids were designed to mimic the conformation of the helical regions of protein A. The conformations of these mimics were examined by a combination of CD and ROE or NOE. In order to further understand the conformation bias toward the real structures of these mimics, some of them are simulated with Quanta/CHARMm through QMD technique.

The Quanta/CHARMm package has shown a poor ability in handling the following operations: drawing big molecules with 2-D sketcher, coloring some portion of a molecule, and exporting the images of an overlay of structures. In the forcefield selection there is no other choice but CHARMm forcefield. In addition, it is impossible for Quanta to be updated indefinitely.

Insight II is a very powerful platform which supports a lot of operations. It can be used to handle all the above problems. Insight II contains many modules inside, such as Builder, Biopolymer, Discover, CHARMm, Docking, NMR_Refines, QuanteMM, and Analysis etc. Furthermore, several forcefields are available for modeling and simulation: cff91.frc, amber.frc, cvff.frc, CHARMM.cfrc, charmm27.cfrc, charmm22.cfrc and charm19.cfrc. So you may do many things what you want to do in this platform, like molecular conformation and property studies. Although both Quanta and Insight II are graphic molecular modeling programs, Quanta only accesses CHARMm simulation engine while Insight II can access CHARMm and Discover engines. So development of QMD on InsightII platform is very significant.

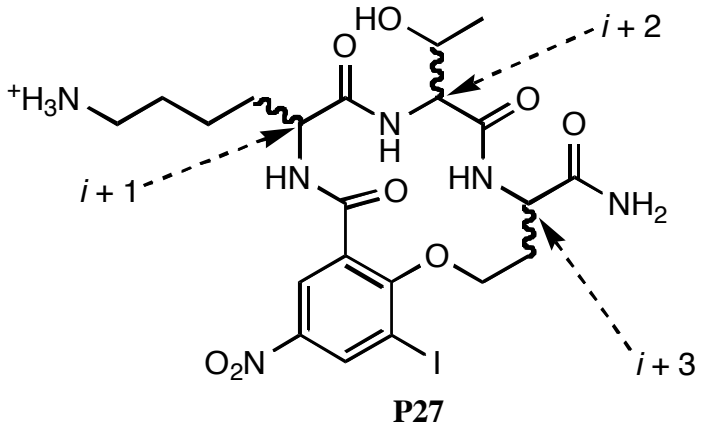
To explore the QMD technique on the Insight II platform, the following molecule **H** (Figure 2.1) made previously in our lab was taken as an example. During the exploration of QMD the greatest challenge is the composition of scripts and the correct linking of these scripts to the main program of Insight II. This is very different from the use of Quanta/CHARMm in which the only thing to do is to click pulldowns.

To obtain some information of the stereochemistry and conformation of a molecule, **P27** was selected for QMD studies (Figure 2.2).



H

Figure 2.1. Structure of a chosen molecule for QMD studies.



P27

configuration

i + 1	*i* + 2	*i* + 3
L	L	L
L	L	D
L	D	L
D	L	L

Figure 2.2. Structure of stereoisomers of **P27**.

2.3 Exploration of the QMD Technique in INSIGHT II/DISCOVER

2.3.1 Molecular Modeling Method on Insight II/Discover

All calculations were performed on a Silicon Graphics O2 workstation with the IRIX 6.3 operating system, using the modeling program Discover_3 (C version, Accelry company) within InsightII (accelry company) environment.

QMD strategy was employed in all simulations; and standard Discover parameters were used. The molecule was modeled as neutral or charged compound. The effects of the solvent was also modeled by a dielectric continuum of □=45, representative of dimethyl sulfoxide (DMSO). The equations of motions were integrated by using Verlet Algorithm with time step of 1 ps. SHAKE was used to constrain all bond length. The starting structure was minimized with 1000 steps of steepest descent (SD) and 3000 steps of the different Newton methods like BFGS, Newton-raphson in order.

The minimized structure was then heated to 1000K by increasing the temperature 10K every 160 time-steps (0.16 ps) from the temperature of 5 K. After being heated, the structure was equilibrated for 12 ps at 1000 K. The molecular simulation run was then performed in the NVT ensemble for a total time of 600 ps. The trajectories were saved every 1 ps. Each of the 600 structures was thoroughly energy minimized using 1000 steps of SD followed by 3000 steps of BFGS or Newton-Raphson.

The structures with lower energies obtained by performing cutoff procedure were selected for further analysis in InsightII. First these structures were clustered into families on the basis of the RMS deviation of the ring backbone atoms. The lowest-energy structure was selected as the representative of this whole study. The lowest-energy structure of each family was selected as the representative of each family. The distances between the protons which interact with one another were measured, and the dihedral angles of the lowest-energy structures of each family were calculated.

The similar molecular modeling method was used in Quanta/ CHARMm.

2.3.2 Script and Procedures with the QMD Technique for Insight II/Discover

A script was written and used for the minimization of a molecule, dynamics, minimization of each of the saved 600 structures, and cut-off. The details are presented in Appendix A.

The detailed procedures of molecular modeling by QMD are presented in Appendix B. They include the methods for building a molecule, minimizing energies, simulating molecules, performing cutoff, and clustering the structures into families.

2.3.3 Results and Discussion for Insight II/Discover

The results from the lowest energy conformer of the first family of each calculation under the following different conditions using InsightII/Discover package were listed in Table 2.2:

(1). Forcefield is cvff, minimization methods are SD and bfgs, and lysine side-chain with NH_2.

(2). Forcefield is cff, minimization methods are SD, newton_raphson, and lysine side-chain with NH_2.

(3). Force field is cff, minimization methods are SD and bfgs, and lysine side-chain with NH_2.

(4). Forcefield is cff, minimization methods are SD and bfgs, and lysine side-chain with NH_3^+.

Table 2.3 also presents the results from the lowest energy conformers of the first family of each calculation under different conditions 5-7 using Quanta/CHARMm package. CHARMm forcefield was employed; and the minimization methods are SD and Adopted Basis Newton-Raphson (ABNR).

(5). Shake on all bonds with hydrogens, and lysine side-chain with NH_2.

(6). No shake was used, and lysine side-chain with NH_3^+.

(7). No shake was used, and lysine side-chain with NH_2.

Even though different conditions were used during the quenched molecular dynamics simulation, all results obtained are not desirable, because no data fit any types of turns using Insight II/Discover package (Table 2.2). Table 2. 3 shows that the results are desirable, and conformations with □-turn types were obtained using Quanta/CHARMm package. There are several factors leading to this: force fields, organic group charge and SHAKING operation.

Figure 2. 3 shows Gaussian energy histograms of 600 structures after minimization. It is very obvious that the energy distributions of 600 structures given by Quanta/CHARMm are closer to Gaussian distribution than those given by InsightII/Discover. The lowest energy structures of each condition except that of condition **7** were displayed in Figure 2.4.

Table 2.2. Summary of QMD Data for Conditions 1-4 (RMSD Threshold 0.7 Å)

items	dihedral angles[a]	different conditions			
		1	2	3	4
E (Glu)	□	-94.02	-130.19	-90.66	-131.80
	□	-71.98	-42.22	-4.45	67.28
K (Lys)	□	-138.57	-120.94	-135.88	68.86
	□	89.84	39.07	72.90	57.24
number in family		89	89	92	91
lowest energy (kcal/mol)		93.027	-65.825	-65.185	-75.088
distance (Å) (CO_i-NH_{i+3})		5.46	4.49	3.89	4.72
types of turn[b]		none	none	none	none

[a]Dihedral angles are defined in page 42. [b]Turn type was determined in term of data from the literature.[81-83]

Table 2.3. Summary of QMD Data for Conditions 5-7 (RMSD Threshold 0.7 Å)

items	dihedral angles	different conditions		
		5	6	7
E (Glu)	□	-70.14	-72.06	-92.80
	□	-29.79	-20.10	47.60
K (Lys)	□	-73.67	-95.40	-165.91
	□	-32.44	7.755	-16.82
number in family		92	95	85
lowest energy (kcal/mol)		-0.4418	-2.2913	-0.0707
distance (Å) (CO_i-NH_{i+3})		2.718	2.252	2.92
types of turn		□III	□I	none

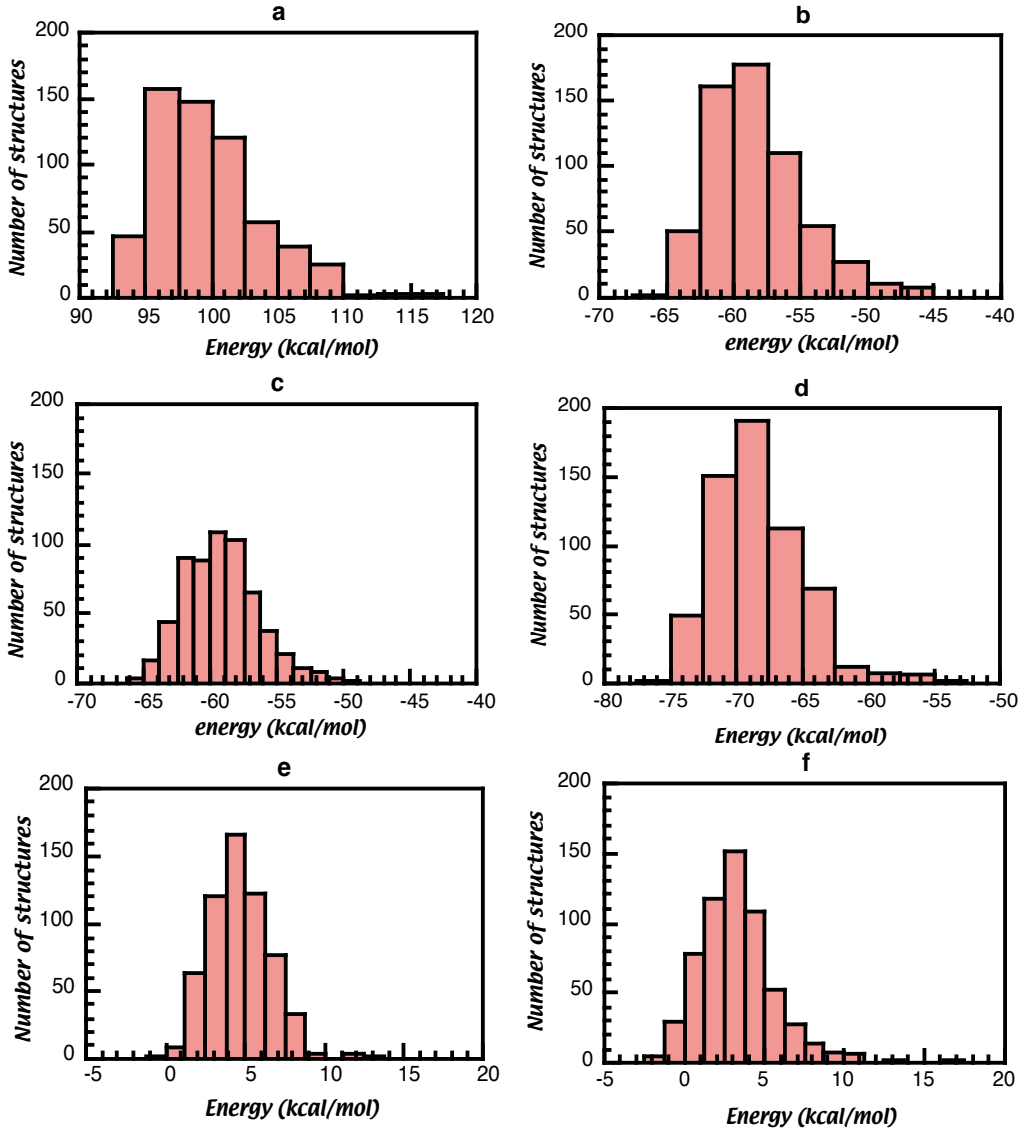


Figure 2.3. a - f Gaussian energy histograms for the QMD studies of compound **H** corresponding to conditions 1 - 6 respectively.

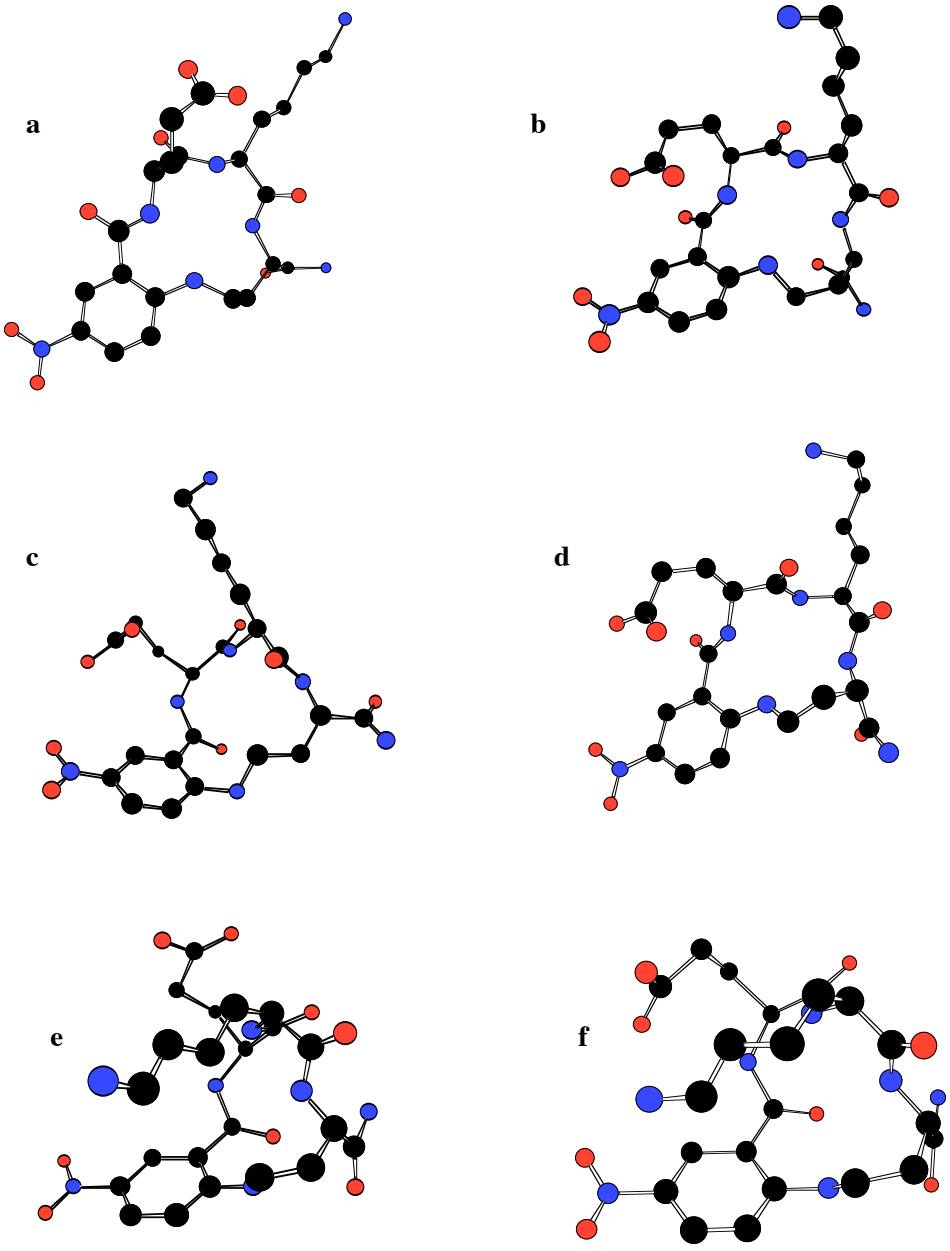


Figure 2.4. a - f The lowest energy structures for the QMD studies of compound **H** corresponding to conditions 1 - 6 respectively.

2.4 Exploration of the QMD Technique in Insight II/CHARMm

2.4.1 Molecular Modeling Method on Insight II/CHARMm

The method is similar to that used for Insight II/Discover. But there are

some differences between these two methods. When the minimized structure was heated to 1000 K, the temperature 10 K was increased every 100 time steps (0.1 ps) from temperature 0 K (not 5 K). The molecular simulation running was performed in either NVE, in which the energy is constant, or NVT, in which the temperature is constant, ensemble for a total time of 600 ps (not only NVT). ABNR was employed to minimize each of the 600 structures instead of BFGS or Newton-Raphson.

2.4.2 Script and Procedures with the QMD Technique for Insight II/CHARMm

A script was written and used for NVE during simulation. When NVT was selected during simulation, the line (TCON TCOUpling 0.400000 TREF 1000.000000) should replace FINALT 1000.000000 in the codes of the simulation section. This script is also used to carry out the minimization, dynamics, minimization of each structure of 600 structures, and cut-off. The details are presented in Appendix C.

The procedures are very simple because Builder was employed to build a molecular structure for InsightII/CHARMm, similar to what it does in InsightII/Discover. The script can be used to handle energy minimization, molecular simulation and cutoff operation. Since it is impossible for InsightII to deal with the separated files including the data from structures or energies. Quanta was then employed for clustering the structures into families and the followed analysis. The procedures are presented in Appendix D.

2.4.3 Results and Discussion for Insight II/CHARMm

Table 2.4 presents the results for the lowest energy conformer of the first family of each calculation in the following conditions using InsightII/CHARMm package (analyzed in Quanta):

(1). Forcefield is CHARMm forcefield, NVE was used during simulation.

(2). Forcefield is CHARMm forcefield, NVT was used during simulation.

(3). Force field is cff, NVE was used during simulation.

(4). Forcefield is CHARMm forcefield, NVE was used during simulation. All of minimization, dynamics, and analysis were carried out in Quanta/CHARMm.

Table 2. 4 shows that all results obtained are desirable because when CHARMm was used as the simulation engine and forcefield, β-turn types can be obtained in either InsightII/CHARMm or Quanta/CHARMm. However, no data fit any turn types using CHARMm as simulation engine and cff as forcefield in InsightII/CHARMm package. This is another example indicating that the forcefield has an important effect on the QMD simulation. Different force fields used in the QMD simulation can cause different results. Gaussian energy histograms of the 600 structures after minimization was presented in Figure 2.5. Figure 2.5 showed the distribution of energy of the 600 structures obtained from Quanta/CHARMm is closer to Gaussian distribution than that obtained from InsightII/CHARMm. The lowest energy structures of each condition are shown in Figure 2.6. The lowest-energy structures of compound **H** in conditions **1** and **4** are similar. Same similarity has been observed for those in conditions **2** and **3**.

Table 2.4. Summary of QMD Data for Conditions 1-4 (RMSD Threshold 0.7 Å)

items	dihedral angles	different conditions			
		1	2	3	4
E (Glu)	φ	-74.33	-69.12	-98.19	-70.14
	ψ	-26.77	-40.44	37.49	-29.79
K (Lys)	φ	-69.22	-66.59	-159.80	-73.67
	ψ	-33.79	-26.80	-23.97	-32.44
number in family		84	69	77	92
lowest energy (kcal/mol)		1.540	-2.508	-69.510	-0.4418
distance (Å) (CO_i-NH_{i+3})		2.759	2.560	3.105	2.718
types of turn		βIII	βIII	none	βIII

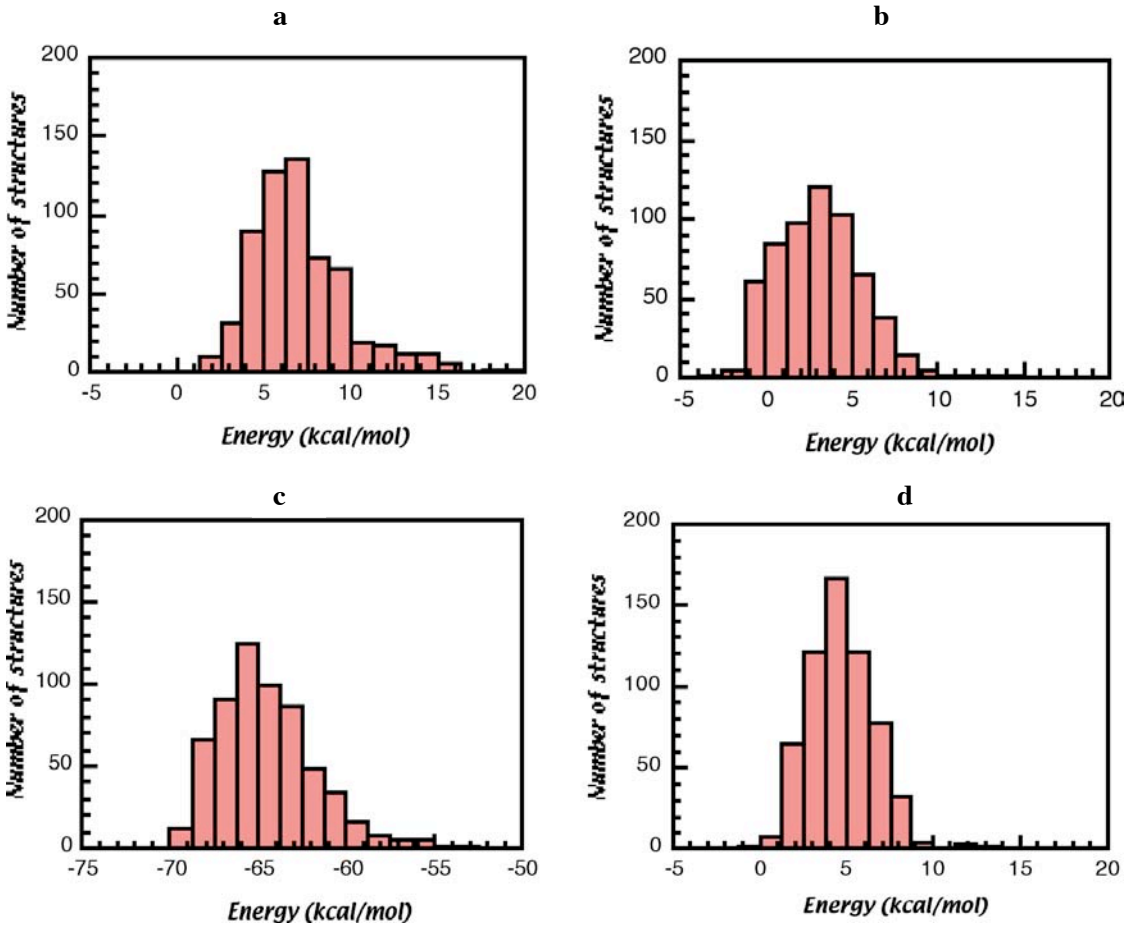


Figure 2.5. a - d Gaussian energy histograms for the QMD studies of compound **H** corresponding to conditions 1 - 4 respectively.

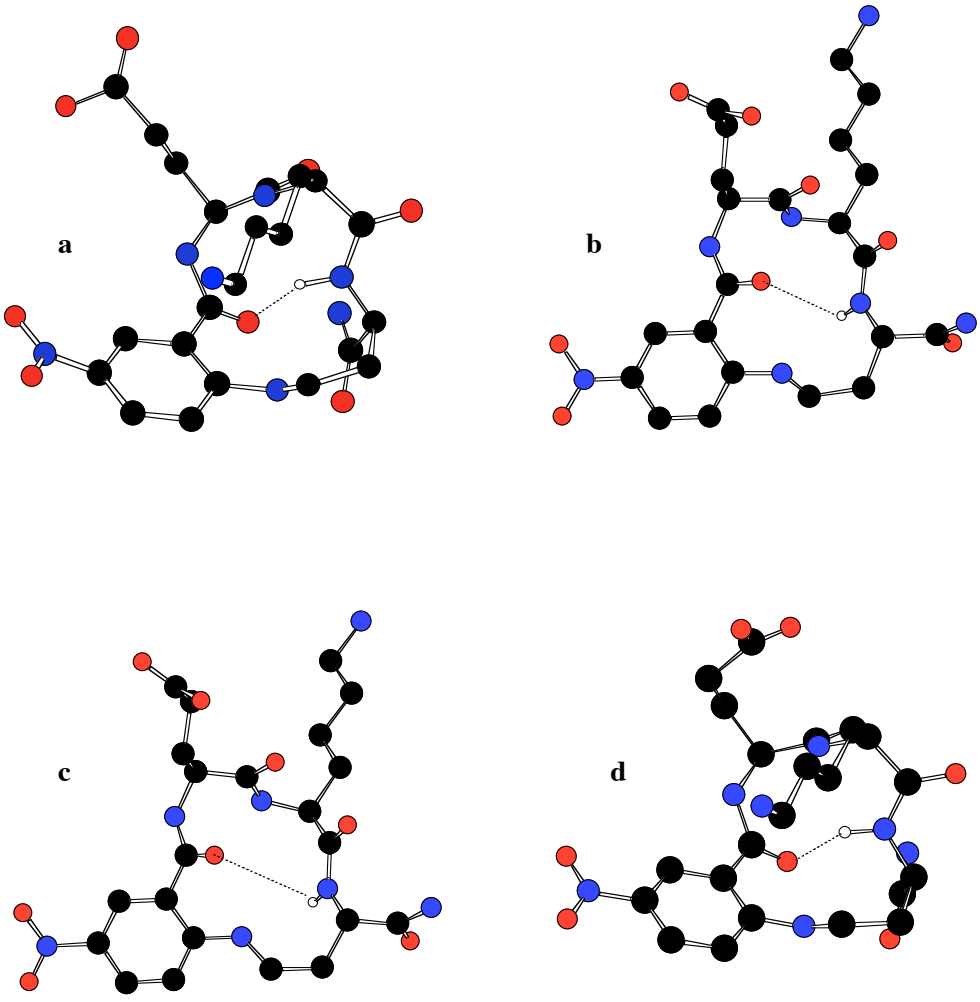


Figure 2.6. a - d The lowest energy structures for the QMD studies of compound **H** corresponding to conditions 1 - 4 respectively.

2.5 QMD Studies of P27 Stereoisomers

2.5.1 Molecular Modeling Method for the Computations of P27

Quanta/CHARMm package was employed for these computations. No SHAKE was used to constrain all bond length with hydrogens. The molecular simulation was performed in the NVE ensemble. Other conditions are the same as those in Insight II/CHARMm.

2.5.2 Results and Discussion for the Computation of P27

Table 2. 5 presents the data results obtained for the lowest energy conformer of the first family of each calculation for four different stereo-structures: LLL, LLD, LDL, and DLL using Quanta/CHARMm package.

All results in Table 2.5 are desirable since each data set of four-different stereo-structures modeling fits one kind of β-turns. There are three types of β-turns for four stereo-structures. Specifically, both LLL and LLD adopt βI turn while LDL and DLL adopt βII and βII' respectively. The lowest-energy structures of these four stereo-structures look different as shown in Figure 2. 7.

Table 2.5. Summary of QMD Data for P27 Stereoisomers (RMSD Threshold 0.7 Å)

items	dihedral angles	stereoisomers of **P27**			
		LLL	LLD	LDL	DLL
K (Lys)	φ	-65.08	-68.90	-72.39	71.77
	ψ	-28.28	-31.14	108.50	-120.20
T (Thr)	φ	-98.83	-82.62	89.36	-94.74
	ψ	22.13	-92.85	35.67	91.89
(HS)	φ	70.12	-75.44	70.96	74.06
	ψ	46.51	-54.50	50.67	62.01
number in family		102	97	85	86
lowest energy (kcal/mol)		1.77	2.16	2.28	1.81
distance (Å) (CO_i-NH_{i+3})		2.245	4.772	2.373	4.418
types of turn		βI	βI	βII	βII'

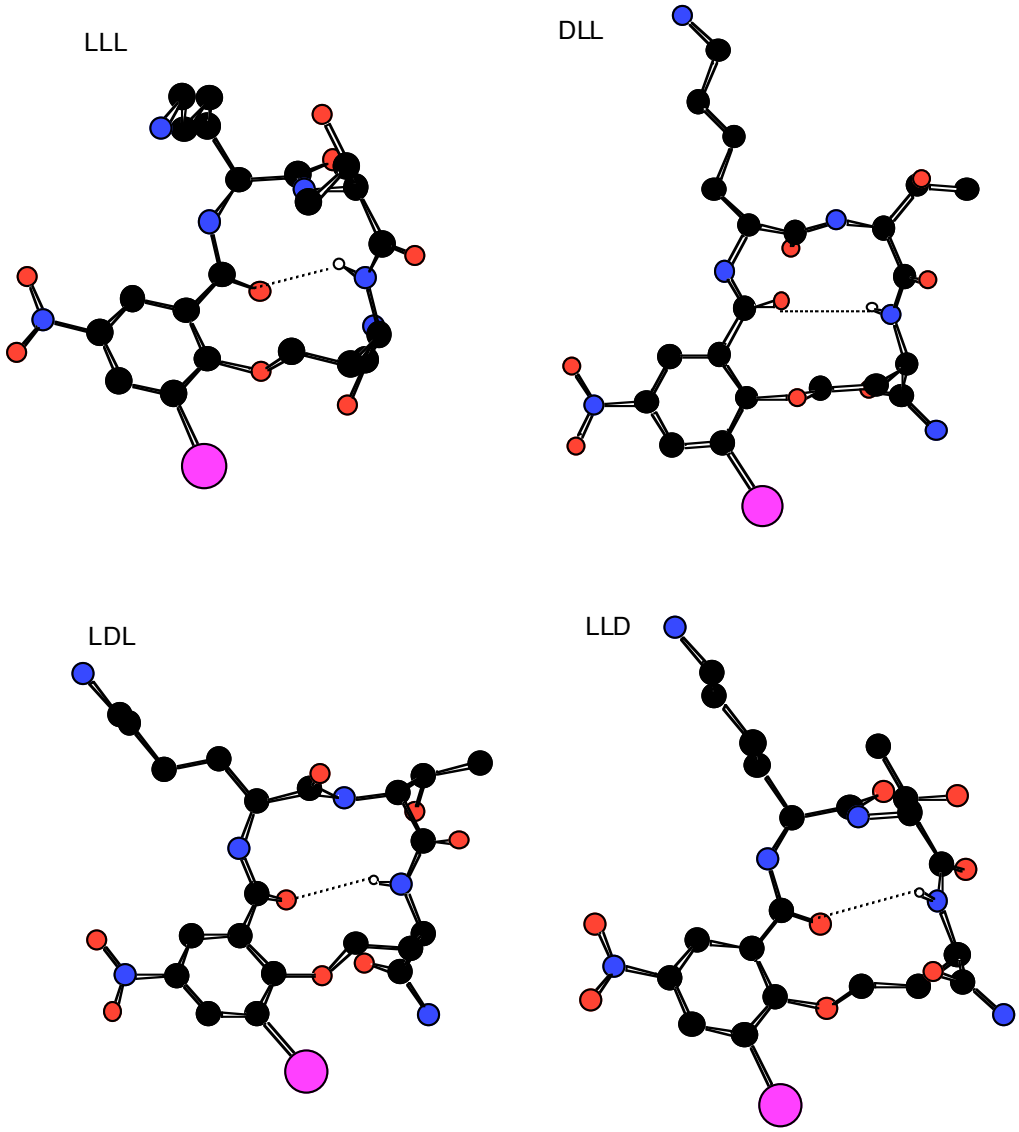


Figure 2.7. The lowest energy structures of four stereoisomers of **P27**.

2.6 Summary

Two written scripts worked very well for QMD study in Insight II platform and succeed in the connection of Insight II program based on the test running. Using the scripts, InsightII/CHARMm may be employed to perform minimization, dynamics (heating, equilibration, simulation), minimization of each of 600 structures, and cut off. Insight II can not be used to carry out clustering families from CHARMm files. In this case, Quanta may be applied. The similar data results obtained from the test run indicate that Insight II/CHARMm may substitute Quanta/CHARMm for QMD studies.

InsightII/Discover may also be applied to perform all of minimization, dynamics, and analysis including clustering of families for QMD studies using the different script. However, the final results are very different from those obtained using CHARMm forcefield and are not desirable according to β-turn types for the selected molecule. This can be attributed to the different forcefields used by Discover and CHARMm.

QMD studies for **P27** suggest that the LLL and LLD isomers prefer a type I β-turn conformation, the LDL isomer prefers a type II β-turn conformation and the DLL isomer prefers a type II′β-turn conformation.

CHAPTER III
PEPTIDOMIMETICS OF PROTEIN A AS POTENTIAL LIGANDS FOR AFFINITY SUPPORTS

3.1 Specific Aims

This project is to design and find peptidomimetics of protein A which resemble its helical regions, have highly selective binding to IgG, and can be produced economically on a large scale, and to test them for bioactivities. Specifically, this research will attempt to:

a. design several peptidomimetics with key residues of protein A and a helix-loop-helix motif constrained by incorporation of Aib;
b. synthesize these peptidomimetics using standard FMOC chemistry;
c. study solution conformations by circular dichroism (CD) spectroscopy studies and NMR spectroscopy studies, such as ROESY and NOESY;
d. simulate the structures of the peptidomimetics in solution using NMR data as limits by QMD;
e. supply samples for testing the bioactivities of the peptidomimetics.

3.2 Background and Significance

3.2.1 Antibodies: Application, Structure and Purification

Antibodies, especially immunoglobulin G, have widespread and varied applications, including ones in diagnostics,[84,85] as bioaffinity ligands in purification of high-value pharmaceuticals (*eg* cytokines and blood-clotting factors)[86] and as probes in diverse biochemical experiments.[85,87-90] Perhaps the most exciting potential applications of antibodies are in the area of therapeutics.[91,92] Animal IgG can be “humanized” so that it is not rejected by the human immune system.[93]

Antibodies are produced by B-cells. However, B-cells themselves can't grow outside an animal. For this reason, Kohler and Milstein[94] developed a theory and method to produce monoclonal antibodies, and were awarded the 1984 Nobel prize for medicine.

The basic unit of most mammalian antibodies is a glycoprotein (~150 KD), which consists of four polypeptide chains, two light chains (~25 KD) and two heavy chains (~50 KD) that are connected by disulfide bonds. Each light chain is composed of two domains, one variable domain (V_L) and the other one constant domain (C_L). There are two types of light chains, Kappa (*k*), lambda (□). In humans, 40% of light chains are □, while 60% are *k*.

Each heavy chain consists of one variable domain (V_H) and three or four constant domains (C_H1, C_H2, C_H3 and C_H4) which depend on antibody isotypes. The hinge region is located in the region between C_H1 and C_H2 domains, and allows flexibility between the two Fab arms of the Y-shaped antibody molecule in order to accommodate binding requirements. The heavy chain also serves to determine antibody classes. Five different classes or isotypes of antibody are distinguished by their heavy chains, that is, IgG to □, IgE to □, IgA to □, IgM to □, IgD to □ (Table 3.1).[91] Each isotype has a unique specialized function, which allows it to perform a certain task in the immune response to different pathogens. IgG1 is the most abundant antibody subclass found in human blood.

The Fc fragment of IgG is composed of a dimer of two C-terminal constant-homology regions of the heavy chain, that is, two constant domains: C_H2 and C_H3. The C_H2 domain consists of about 110 amino acids, while the C_H3 domain consists of about 106 amino acids. The two monomers are related with a perfect two-fold axis. Figure 3.1 shows the crystal structure (PDB: 1MCO)[95] and schematic representation of antibody IgG.

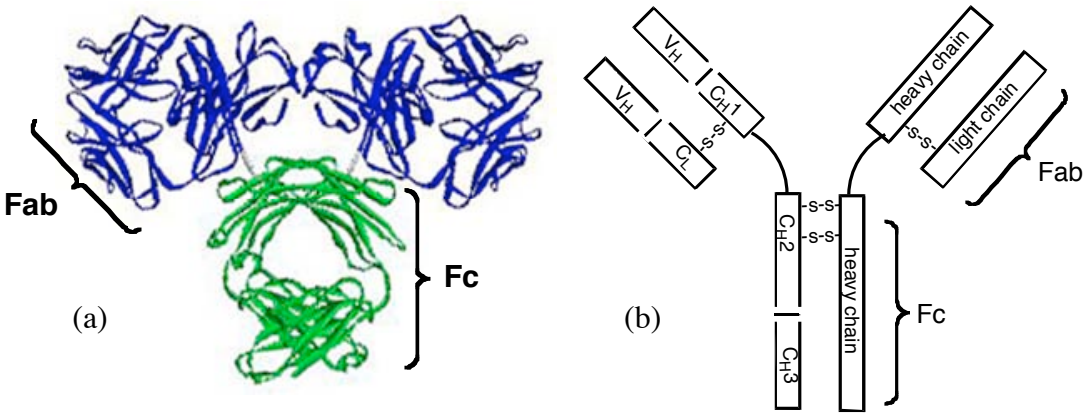


Figure 3.1. Antibody. (a) crystal structure of intact human IgG1 which lacks functional hinge regions, (b) schematic representation of an antibody.

Table 3.1. Characteristics of Human Antibody Isotypes

antibody class	IgG				IgE	IgA		IgM	IgD
antibody subclass	IgG1	IgG2	IgG3	IgG4		IgA1	IgA2		
heavy chain	γ_1	γ_2	γ_3	γ_4	ε	α_1	α_2	μ	δ
molecular weight (kDa)	146	146	165	146	188	160	160	970	184
adult serum level (mg/ml)	9	3	1	0.5	5.00E-05	3	0.5	1.5	0.03
serum half-life (days)	21	20	7	21	2	6	6	10	3

IgG is the most important class of antibodies and obtained from human plasma, ascites fluid, fetal calf serum, culture supernatants of engineered cells (*eg* hybridoma and bacterial cells) and other sources, then it must be purified for most applications. However, the purity of IgGs is critical, particularly for therapeutic applications. Usually its purification is achieved via a chromatographic technique, and there are several options for this. Size-exclusion and ion-exchange[96,97] methods have both been used, but affinity chromatography is the most widely applied.[86,98,99] Monoclonal antibodies (mAbs) designed to bind a specific antigen may be purified using supports based on that antigen.[100] A drawback in that approach is that the binding constants for the antigen tend

to be so high that harsh elution conditions are required and denaturation of the IgGs can occur. Moreover, for widespread large-scale purification of mAbs, methods that work irrespective of the mAb epitope are most important. Such generalized methods for affinity chromatographic purification of antibodies rely on their interactions with ligands that bind non-variable regions on the antibody surface.[98,101] Usually these ligands are naturally occurring proteins that have this characteristic, like protein A from *Staphylococcus aureus* and protein G from *Streptococcus*.[102,103] Affinity columns based on protein A are the most widely used for several reasons. First, protein A is well-characterized and may be obtained from recombinant bacteria.[102,104] Second, protein A interacts with the Fc fragment of IgG with a high affinity constant (~10^{-7}M).[105] Third, protein A is stable over a wide range of pH 2-11, it can refold after treatment with denaturating solutions like urea and guanidinium salts,[106] it can even be cleaned using 0.5 M NaOH, hence the affinity supports can sometimes be recycled.[107] Finally, protein A is easily coupled to supports.

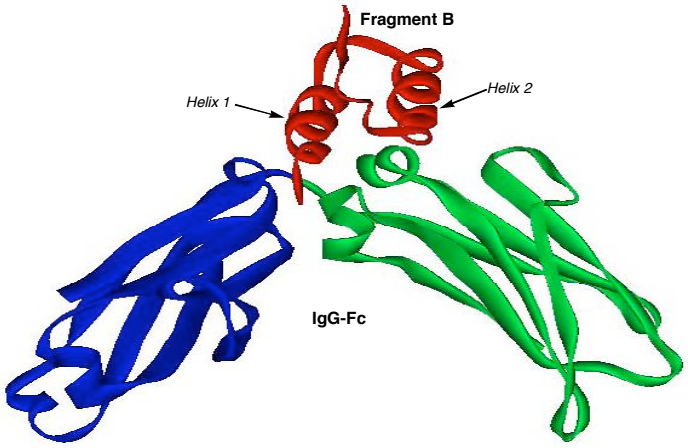
3.2.2 Protein A and Its Binding with Immunoglobulin G

Protein A is a cell wall component of *Staphylococcus aureus*. It has five highly homologous IgG-binding domains A, B, C, D and E, each of which consists of 58-62 amino acid residues.[108,109] The structure of fragment B of protein A has been determined by X-ray crystallography and NMR spectroscopy.[104,110,111] A major part of the structure of fragment B is composed of two α-helices while the rest is folded irregularly (Figure 3.2c). The residues from these two helices in protein A participate in the protein-protein interface that is located in the hinge region that connects the second and the third constant domains of the heavy chain (C_H2 and C_H3) of Fc.

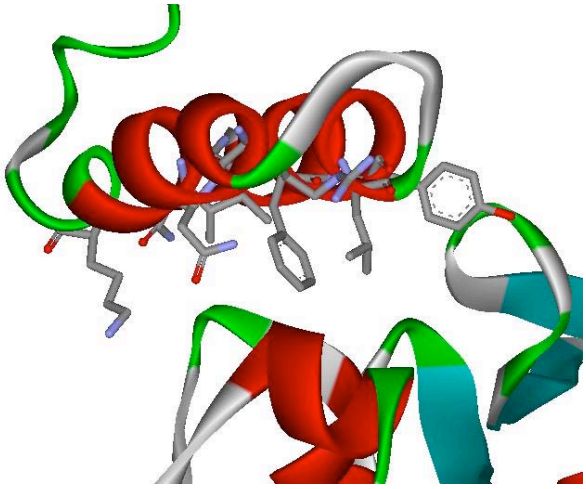
Generally, a protein contains a number of domains which are considered to be the smallest functional elements of protein structure, and can create great diversity for protein-protein interactions.[112] Protein A recognizes antibodies by binding to their Fab or Fc fragments. The binding of protein A to the Fc of the antibody is very strong and well known, while the behavior of binding between protein A and the Fab of IgG is rather less well characterized.[113] However, protein A shows weak binding activity with the Fab of IgE, IgA and IgM.[114-116]

Crystallographic data for the complex between the key helix-loop-helix region of protein A (the "B-domain") and the human IgG Fc fragment (protein databank numbers:1l6x and 1FC2; Figure 3.2a) is thought to illustrate the key hot-spots for the interaction.[104] It features contact of residues in the two helices of protein A/fragment B (residues Gln128-Leu136, Glu144-Asp155) with β-turns of the IgG Fc region. This structural model is also supported by data from NMR[117,118] and mutational studies.[119] Consequently, it has been proposed[120] that the following protein A residues are important for binding the Fc region of IgG: Phe124, Phe132, Tyr133, Leu136, Ile150, Lys154, Gln128, Asn130, Asn147, and Arg146. Of these, Phe132 and Tyr133 may be pivotal; they form a hydrophobic pocket that encapsulates the Ile253 of IgG (Figure 3.2c).

(a)



(b)



(c)

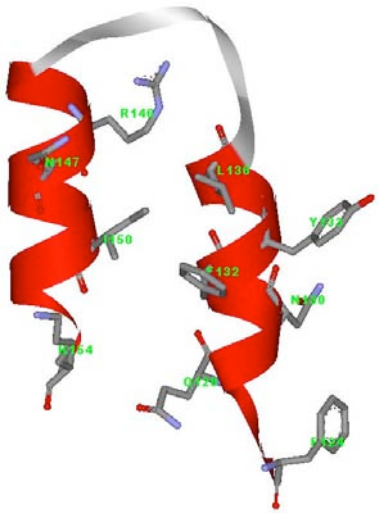


Figure 3.2. (a) The B domain of protein A complexed with Fc fragment of IgG; (b) an expanded view of the same interaction from a different perspective; (c) "hot-spots" from protein A involved in the Fc binding.

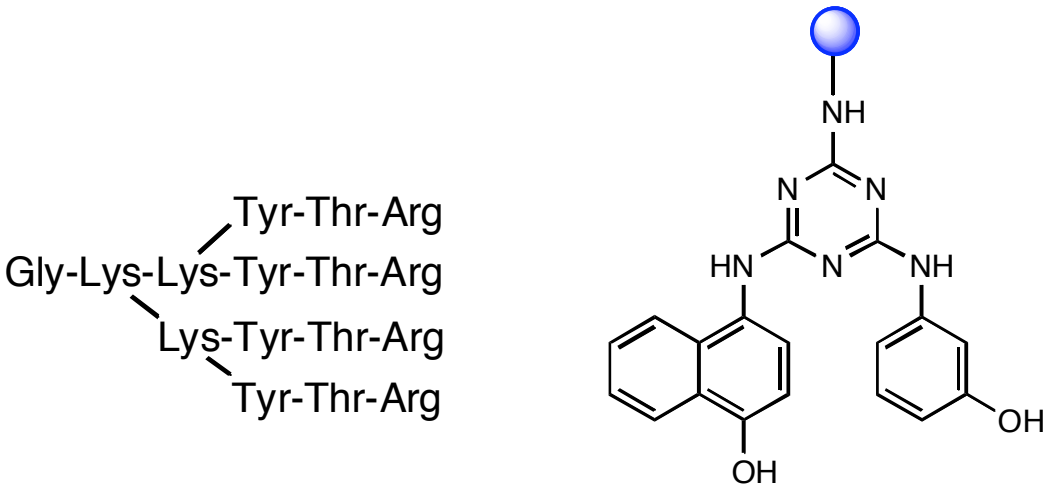
3.2.3 Structure Based Design of Protein A Mimics

Even though protein A binds IgG with high capacity and selectivity, several deficiencies exist with current purification strategies, like high cost, especially for large-scale applications. Protein A is an expensive bacterial product, and its production includes expensive and laborious procedures. Another drawback is that protein A may leach from the matrix and contaminate the isolated immunoglobulins.[121,122] For these reasons, inexpensive, synthetic ligands with appropriate affinities for the non-variable regions of IgG are extremely interesting as potential ligands for affinity supports.

As early as 1987, Proath[123] found that artificially thiophilic ligands might be used to replace protein A in the affinity purification of antibodies. Later Elkak[124] proposed hystidyl ligands to be used; Khatter[125,126] suggested that supported, functionalized, pyridyl-pyridinium salts could also be used. However their poor selectivity for antibodies or incompletely characterized structures limited their application. Recently, with the development of computer-aided molecular design and the utilization of combinatorial technologies,[127] Fassina's group[128] designed and prepared a multimeric peptide library which contains the general formula $(X1\text{-}X2\text{-}X3)_4$-K2-K-G. After screening the activity of three sublibraries, they found that the lead compound is TG19318 (Figure 3.3). TG19318 is a very active inhibitor of the binding of rabbit IgG to protein A, and has broader specificity than protein A. TG19318 may be used to purify not only IgG from different sources, but also IgA, IgE, IgM and IgY.[129-131] The apparent affinity constant for TG19318 binding to IgG was found to be $k_a = 3 * 10^5\ M^{-1}$ and $k_d = 0.3\ \mu M$.[132] The purities of isolated immunoglobulins were judged using SDS-PAGE and ELISA. The optimal binding conditions between TG19318 and immunoglobulins are at pH 6.5-7.5 with buffers at low ionic strength.[133] However, there exist some limitations for TG19318 use. TG19318 has poor selectivity, and does not distinguish among different immunoglobulin isotypes. It may be used to purify monoclonal antibodies, but not to isolate either polyclonal IgA or polyclonal IgM from serum. The binding mechanism between TG19318 and various immunoglobulins is unknown. There is no direct evidence about the binding between TG19318 and the Fc portion of Igs.

Lowe's group[120] designed, synthesized and screened a nonpeptidic library using triazine as a scaffold to mimic a key, Phe132-Tyr133, dipeptide motif on fragment B of protein A that is believed to play an important role in the interaction with the Fc fragment of IgG.[104] They found a completely different kind of lead structure, termed as 22/8. This ligand prepared using solution phase has a lower affinity constant: $k_a = 1.4 * 10^5 M^{-1}$ and $k_d = 7.1 \mu M$ for the binding of human IgG, than that of protein A, while, attached on agarose matrix, 22/8 (Figure 3.3) exhibits a high binding capacity of IgG: 151.9 mg IgG/g moist wet gel, probably due to higher ligand concentration in the affinity matrix.[134] Like TG19318, the ligand 22/8 displays a broader specificity, and can bind immunoglobulins from different species, the binding order being as follows: human>chicken>cow>rabbit>pig>horse>rat>goat>sheep>mouse. Its binding pattern is not identical to that of protein A. However, the binding mechanism between 22/8 and IgG is also unknown. The purity of the isolated IgG was judged using only SDS-PAGE. There are no data about ligand leaching and toxicity in the literature.

Nonetheless these lead ligands show some binding to IgG, and their application is a great help for large-scale purification because of their stability toward cleaning-in-place procedures, they might not bind IgG with high selectivity, capacity and appropriate affinities. None of these have found widespread commercial application, so we have begun to explore peptidomimetics of protein A as potential ligands for affinity supports.



TG19318 **22/8**

Figure 3.3. Structures of protein A mimics.

3.2.4 Techniques for the Binding Study of Ligands with IgG

3.2.4.1 Affinity Chromatography

Affinity chromatography is a powerful fractionation technique for large-scale purification of biotechnological products. Antibody purification is carried out using natural affinity ligands (such as proteins A and G) via this technique. Affinity chromatography is composed of five steps in the following order: activation of the matrix, coupling of ligands, adsorption of the antibodies, elution, and regeneration of the affinity matrix. The theory and application of affinity chromatography have been depicted in books and articles.[135-137]

The principle of affinity chromatography (Figure 3.4) is based on molecular recognition between antibodies and ligands. Commonly, any antibody that needs to purify has an inherent binding site called recognition site. The antibody is able to recognize a ligand molecule via this site. If the ligand is immobilized on a polymeric matrix, when a biological sample, such as blood serum, is passed through the affinity column, the immobilized ligand can capture the antibody selectively. The elution buffer is passed through the column by changing external conditions, such as pH value, ionic strength, solvents and temperature, so that the complex consisting of the

antibody and ligand is no longer stable; the antibody can then be released from the complex and eluted in a purified form.

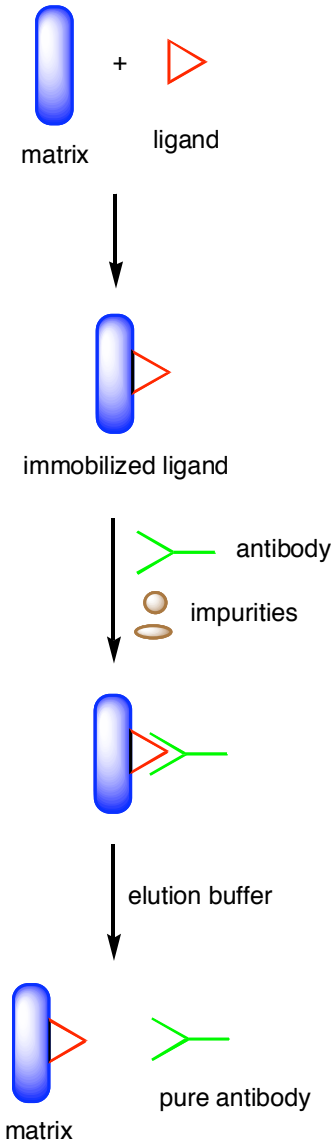


Figure 3.4. Schematic presentation of antibody purification by affinity chromatography.

3.2.4.2 Saturation Transfer Difference NMR Spectroscopy

Various NMR based techniques, for example, transferred NOE, diffusion or relaxation-edited NMR, ^{15}N chemical shift studies and saturation transfer difference (STD), have been employed to facilitate the screening of compound mixtures for components which bind protein targets.[138-143] STD NMR spectroscopy is a very useful and important technique. STD, which means the difference between a saturation transfer spectrum and a normal NMR spectrum, can provide a fast and sensitive method to screen library compounds for their binding activities to the protein, and to identify the binding epitope of ligands to the protein because the ligand residues which directly contact to the protein exhibit much stronger STD signals than others. The range ($10^{-3} \sim 10^{-8}$M) of dissociation constants for ligands to bind to IgG is required.[144] Figure 3.5 shows the effect of a selective saturation pulse on the antibody to bind ligand molecules illustrated as an ellipse. First, several antibody resonances are saturated by the selective pulse. This saturation is then spread over the entire antibody as the intramolecular saturation transfers. This change is indicated by the shading of the antibody, from light shade to deep shade. Even though the resonances of the small ligands are not directly affected by the selective pulse, those ligand molecules which interact with the antibody are saturated as intermolecular saturation transfers. The saturated ligands are released from the complex consisting of ligands and antibody, and enter the solution via chemical exchange. The released ligands may be detected. Some STD NMR studies for my peptide ligands to bind IgG were performed, but the results of these studies could not be validated since I do not have standard ligands, which are already known to bind to IgG, for comparison.

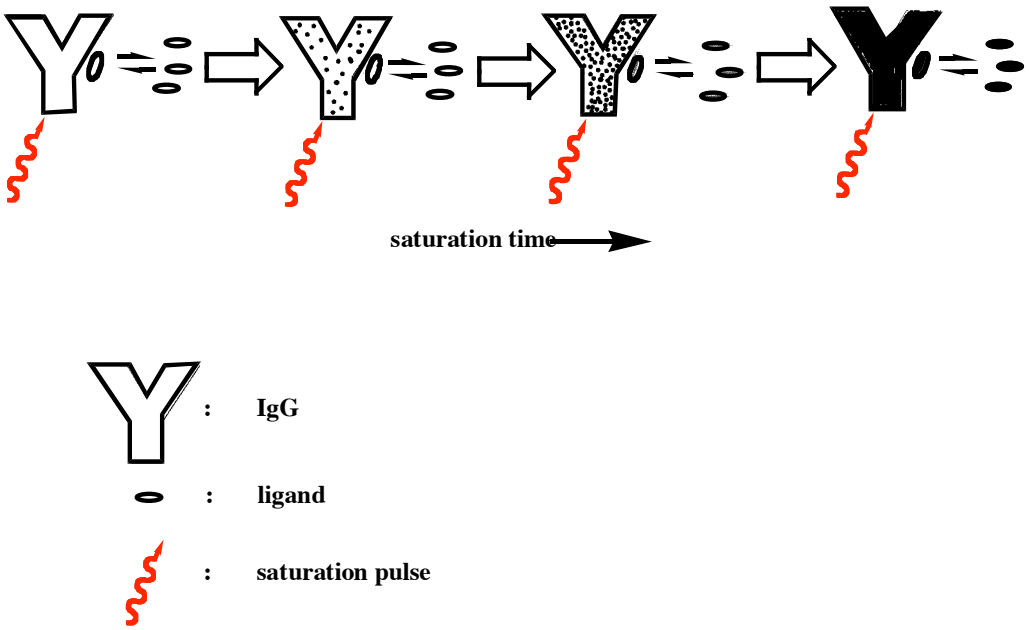


Figure 3.5. Basic principle of STD NMR spectroscopy.

3.2.5 Conformational Studies of Mimetic Ligands

Conformational studies of mimetic ligands are very necessary in biological and medicinal chemistry. They reveal useful information concerning the preferred conformation of the mimetic ligand in solution and the important residues in the preferred conformation. They also provide information on ligand-receptor interactions, which is helpful to revise ligand design. The conformational studies presented here include NMR, CD and molecular modeling studies.

3.2.5.1 NMR Studies

A combination of several NMR techniques is a very powerful tool in practice for structural determination of peptidomimetics, of which the quality crystals are not easily accessible.[145,146] Useful information for identification of secondary structures, such

as ^{1}H-^{1}H distance, coupling constant ^{3}J, torsion angle, etc., can be obtained. Table 3.2 summarizes several NMR techniques in this study.

Table 3.2. Important NMR Techniques Used in This Study

NMR technique	observed data	useful information
1D ^{1}H NMR	chemical shift and ^{3}J-coupling constant	torsion angle
DQF-COSY and TOCSY	^{3}J-coupling crosspeaks	through-bond connectivities residue assignment
ROESY or NOESY	proton close contacts	sequential assignment distance geometry

^{3}J-Coupling constant. It is an important parameter in the determination of peptidomimetics structures,[147,148] and directly relates to the dihedral angle or torsion angle, that is, the spatial arrangement of residues. The dihedral angle is often used to determine protein secondary structure. C-N-Cα-C *in para* are the atoms along a protein backbone. The dihedral angle of (C, N, Cα, C) is called ϕ torsion angle while the dihedral angle of (N, Cα, C, N) is called ψ torsion angle. In terms of the following Karplus equation, the torsion angle (θ) values in a polypepide may be estimated.[149]

$$^{3}J_{HN\alpha} = 6.4 \cos^{2}\theta - 1.4\cos\theta + 1.9$$

where $\theta = | \theta - 60^{o} |$

For ideal α-helix and β-sheet, the estimated values[147,150] of torsion angle and coupling constant $^{3}J_{HN\alpha}$ are:

For α-helix, $\phi = \sim -60^{o}$ and $^{3}J_{HN\alpha} < 6$ Hz

β-sheet, $\phi > -118^{o}$ and $^{3}J_{HN\alpha} > 8$ Hz.

NOE or ROE-Observable Proton Close Contacts. ^{1}H-^{1}H distances can be estimated by either the nuclear Overhauser effect spectroscopy (NOESY)[151-153] or the rotating frame Overhauser effect spectroscopy (ROESY)[154,155] experiments and measuring cross peak intensity. Generally, NOE includes three effects: negative, null and positive enhancements. Negative enhancement is observed for large molecules whose molecular weight is greater than 2000, while positive enhancement is expected for small molecules whose molecular weight is less than 1000.[155,156] If medium-sized molecules are used for NOESY experiment, the NOE is often close to zero. In this case, ROESY pulse sequence should be chosen because a weak radiofrequency and various mixing times are applied, and this minimizes false NOE effects and HOHAHA cross peaks and then increases the cross-peak intensity with correlation time.[157] Overall, NOESY is only useful for proton close contact information of small and large molecules, while ROESY is applicable to all-sized molecules, especially to medium-sized molecules.

NMR Characteristics of Ideal Helix. helix is one of the most important secondary structural motifs in proteins.[158,159] It often plays a key role in molecular recognition. In a helical structure the polypeptide backbone is tightly wound around the long axis of the molecule, and the *R* groups of the amino acid residues protrude outward from the helical backbone. The repeating unit is a single turn of the helix, which includes 3.6 amino acids. In case of a standard helix, the three-dimensional path through its axis is a perfectly linear. However, helices have been observed to be non-linear due to solvent induced distortions,[160] peptide bond distortions[161,162] and others.[163]

NOE-observable ^{1}H-^{1}H distances and the coupling constant ($^{3}J_{HN\square}$) are two important parameters in the determination of helix structure of polypeptide. The distance between the hydrogen atoms X and Y located in the amino acid residues in the sequence positions *i* and *j* respectively, is presented by $d_{XY}(i, j)$. Specifically, $d_{\square N}(i, j) \equiv d_{\square N}(\square H_i, NH_j)$, $d_{NN}(i, j) \equiv d_{NN}(NH_i, NH_j)$, etc. Figure 3.6 shows some distances defined in the direction from the N to C terminus of the polypeptide chain.[164]

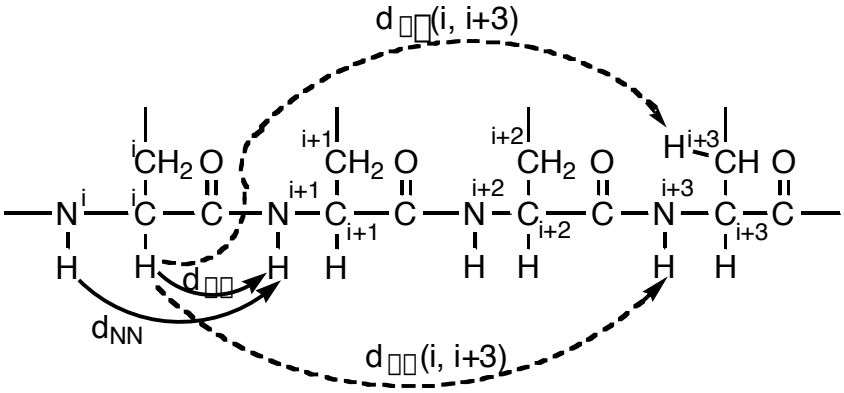


Figure 3.6. Some short-range proton-proton distances in a peptide chain.

NOEs may be classified as strong, medium and weak, which are represented by the thickness of the line. The characteristic patterns of short-range NOEs in an ideal helix are described in Figure 3.7.[165]

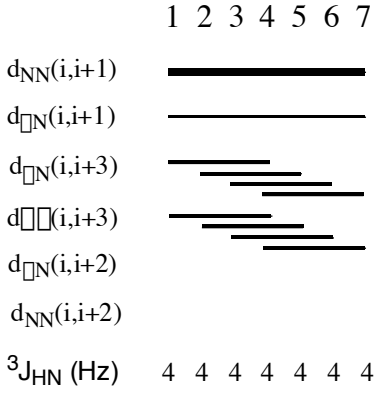


Figure 3.7. Short-range NOEs observed in an ideal helical conformation.

3.2.5.2 Circular Dichroism Studies

Circular dichroism (CD) is often applied for the structural characterization of peptides or proteins since CD spectra are remarkably sensitive to the backbone conformation of proteins.[166,167] CD spectra of peptides or proteins in the 190-230 nm region result from the peptide amide chromophores, and the chiral environment of all amide bonds and amide-amide interactions.[168]

Different secondary structural types of proteins display distinct CD characteristics. Helical structures are characterized by a strong positive π-π^* band between 190-195 nm, a negative π-π^* minimum at 208 nm and a negative n-π^* minimum at 222 nm.[158,169] A strong positive π-π^* band between 195-200 nm and a negative n-π^* band between 215-220 nm are the characteristics of β-sheets.[170,171] For random coils, the CD spectra often show a strong negative π-π^* minimum at 200 nm and a weak positive n-π^* band near 220 nm.[172]

The CD spectrum of a protein often reflects the secondary structural content of this protein. The CD signal at 222 nm is used for a rough estimation of the helical content via some equations.[173]

A number of solvent systems may be used to promote the helix characteristics of peptides, such as trifluoroethanol (TFE)[174,175] and hexafluoroisopropanol (HFIP).[176] The main effects of TFE and HFIP result from their significantly weaker basicity.[177] In these solvent systems, hydrogen bonding of amide protons to the solvents is decreased, and intramolecular hydrogen bonds are strengthened. Consequently, this effect stabilizes secondary structures. In addition, TFE is a less polar solvent, and interrupts hydrophobic interactions and denatures the tertiary and quaternary structure of proteins.[178]

3.2.5.3 Molecular Modeling Studies

Molecular dynamic simulation may be applied to predict various possible conformations for particular molecules and to visualize the structures of the conformations. Generally, two types of simulations are used for the modeling of peptides or mimetics. One is NMR data independent simulation wherein the NOEs or ROEs

obtained from the 2D NMR experiments are not required for input as restraints for the simulation. Only the parameters from the software are used. The other is NMR data constraint simulation wherein the NOEs or ROEs data must be set up as limits before simulation runs. Conformations will then be generated under this condition using the selected simulation methods.

3.3 Design of New Protein A Peptidomimetics

Protein A includes two helices, it is very important to understand helical mimics. Commonly, helical conformation is not the minimum energy conformation of short peptides which consist of less than 20 amino acids. A variety of approaches have been adopted to stabilize helical peptides, which include: first, noncovalent side chain constraints, such as hydrophobic interactions,[179] metal ion induction,[180-182] and salt bridges;[183,184] second, covalent bonding between side-chains, such as i to i + 3, i to i + 4, or i to i + 7 linkages,[185-189] and disulfide linkages which span non-adjacent loops between i to i + 4 and i to i + 7;[190-192] third, incorporation of unnatural amino acids into short peptide sequences (for example, □-aminoisobutyric acid (Aib) is able to stabilize helical structures in short peptides).[193-197] Besides, short peptides may be stabilized by some solvents, such as trifluoroethanol[198] and hexafluoroisopropanol,[176] intrinsic helix dipole,[176,199] and *C*- or *N*- terminal capping motifs.[200-202] Based on the helix-loop-helix structure of protein A and the approaches summarized above, the following protein A peptidomimetics are proposed (Figure 3.8). Of seven peptidomimetics, five include two Aibs and resemble Balaram's helix-turn-helix motif.[194] Aminocaprionic acid [Acp, $NH_2(CH_2)_5COOH$] is used as a linker similar to the loop of protein A.

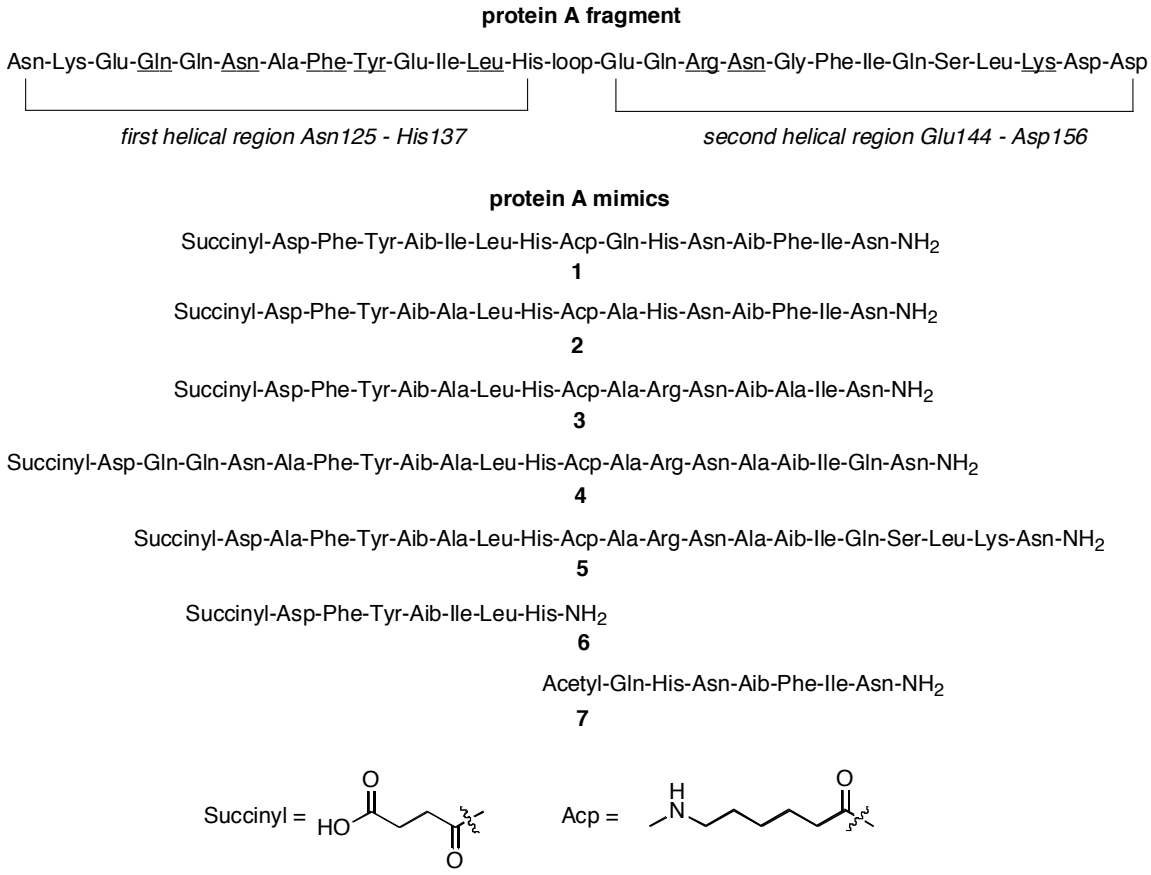


Figure 3.8. Sequence of the helix-loop-helix fragment of protein A and sequences of the peptidomimetics targeted.

Compared with the protein A original sequence, peptidomimetics **1-5** contain most of the amino acids identified as " hot-spots", which are underlined, to bind IgG/Fc. A key, Phe132-Tyr133, dipeptide motif is involved in each of peptidomimetics **1-6**. However, all of the peptidomimetics are much shorter than the protein A original sequence, meaning that they are much easier to make. In peptidomimetics **1-6** the *N*-terminus is capped with succinic acid which provides a negative charge to stabilize the intrinsic helix dipole. The Asp next to succinyl reinforces this stabilization. Succinyl-Asp is used to form a *N*-terminal capping box.[199,200] The Aib is incorporated to the peptidomimetics and expected to enhance helix stability. The Asn used as a *C*-terminal capping box is intended to stabilize the helical conformation by formation of a side-

chain to main-chain i + i-4 interaction.[201] Replacement of Ile, Phe or Gln with Ala in the peptidomimetics **2** □ **5** is intended to enhance helix stability.[203] In peptidomimetics **4** and **5** the *N*-terminal or *C*-terminal side is expanded to see if this is important to both helix stability and binding with IgG. Peptidomimetics **6** and **7** were minimalist designs to mimic only the first and second helical regions in the protein A helix-loop-helix motif, respectively.

3.4 Solid Phase Syntheses of Peptidomimetics Using Fmoc Approach

All the peptide sequences were prepared using Rink's amide linker on HypoGel 400 Ram resin via the conventional Fmoc approach.[204-206] Peptidomimetic **1** was taken as an example, and its synthesis by solid phase was presented in Scheme 3.1. In this scheme there are many repeated steps (coupling between –COOH and $–NH_2$, and removal of Fmoc). Most of the couplings were performed using HBTU/HOBt[207] except that of Aib. Since Aib is a spatially hindered residue the coupling with Aib is very difficult and does not work well using these coupling reagents. Two other coupling methods with several different conditions were tried. One is where Fmoc-Aib-OH was first transformed to Fmoc-Aib-F using cyanuric fluoride,[208] and the Fmoc-Aib-F was then used to react with other amino acids. The other is where tetramethyl fluoro formamidinium hexafluorophosphate (TFFH)[209] made by ourselves was used as coupling reagent instead of HBTU/HOBt during the coupling reaction. Both methods work well, and the second one was selected for the syntheses of seven peptidomimetics because several grams of TFFH may be prepared at one time with high yield. In order for Aib to be coupled to the sequence successfully, double coupling cycles[210] are required, and each used 4 equivalents of Fmoc-Aib-OH, 4 equivalents of TFFH and 8 equivalents of DIEA. For the same reason, the coupling reaction of the amino acid following the Aib coupling needs to carry out twice too. Table 3.3 summarizes purity, yield and MS (MALDI) data for seven peptidomimetics.

Scheme 3.1. Solid Phase Synthesis of Peptidomimetic 1

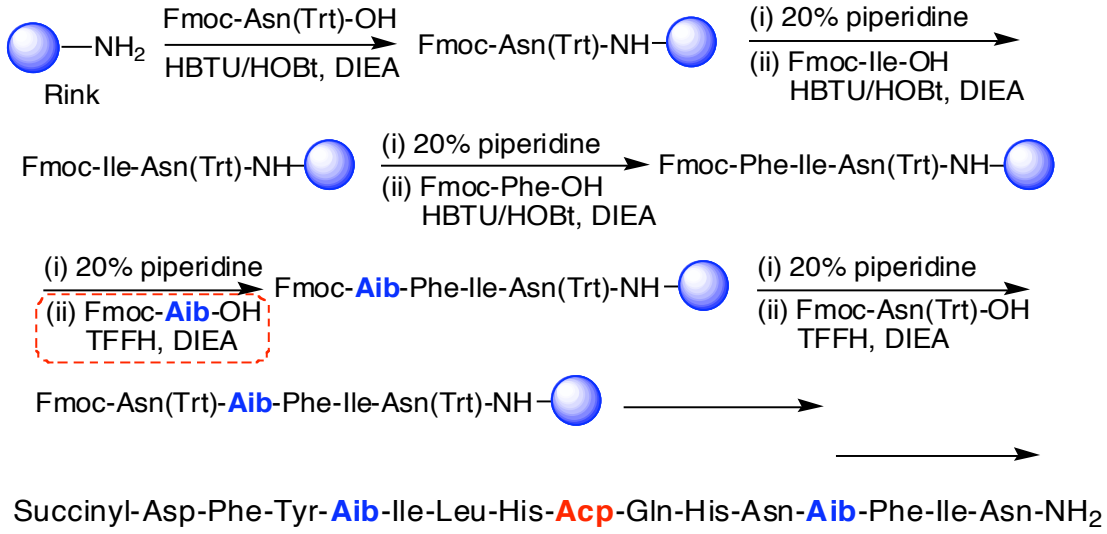


Table 3.3. Summary of Purity, Yield and MALDI-MS Data for 7 Peptidomimetics[a]

compound	purity[b] (%)		yield	MALDI MS for (M^++H)	
	UV	Sedex	(%)	calcd	found
1	91	92	90	1944.2	1944.6
2	66	74	78	1845.1	1845.6
3	81	89	82	1788	1788.4
4	87	90	88	2357.6	2357.5
5	68	73	80	2315.6	2315.4
6	93	95	95	992.1	992.2
7	78	81	72	899	899.3

[a]HypoGel 400 RAM (0.53 mmol/g) was used for the syntheses of the seven peptidomimetics. [b]Purity assessed by HPLC for crude via monitoring UV absorption at 254 nm and using an evaporative light scattering detector (Sedex).

3.5 Conformational Analyses of Peptidomimetics

The strategy used in this study was first to record CD spectra for all the peptides in pH 4 phosphate buffer (20 mM), then with varying concentrations of 1,1,1,3,3,3-hexafluoropropan-2-ol (HFIP). Recording CD spectra using buffer medium revealed if the peptide sequences showed any biases towards helical conformations that were observable by that technique. Similar experiments using HFIP/buffer were used to reveal the presence or absence of helicity under conditions that are more conducive to this. Consequently, these analyses were used to select the compounds that would be studied more thoroughly by NMR and molecular simulations.

3.5.1 CD Studies of Peptidomimetics

The CD spectra of peptidomimetics **1**, **3**-**7** (50 μM) in phosphate buffer (20μM, pH=4) and at different concentrations of HFIP are shown in Figures 3.9, 3.10 and 3.11. Peptidomimetic **7** shows no observable bias towards helical conformations under any conditions, and its CD spectra are characteristic for random coil conformations. The spectra recorded in water indicate largely unstructured fragments with a small contribution of α-helical secondary structure for peptidomimetics except for **6**. The addition of HFIP stabilized an α-helical conformation. There are one distinct maximum between 190-195 nm and two minima at 222 nm and 208 nm, which are the characteristics of α-helical secondary structure. Initially, the helicity increases steadily with the concentration of $(CF_3)_2CHOH$. The helicity does not change further or changes little when the concentration of $(CF_3)_2CHOH$ reaches some values, for example, 20% (by volume) for peptidomimetic **1** and 15% for peptidomimetics **3**-**6**, which were selected for NMR studies.

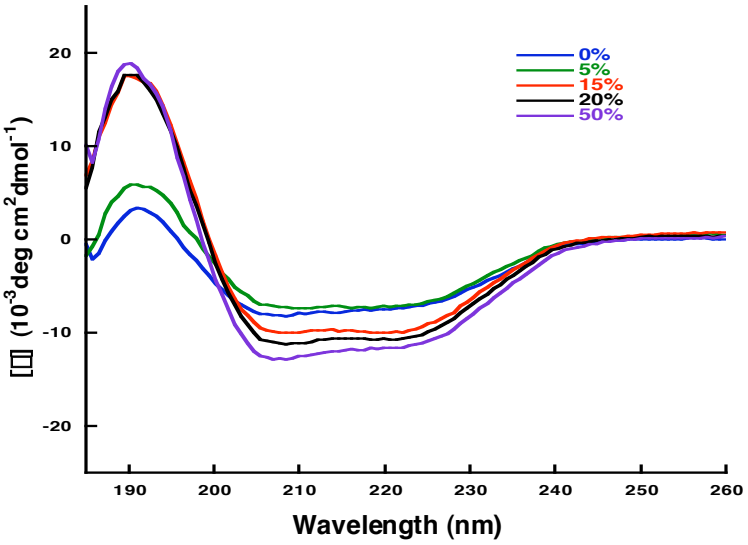
The CD signal at 222 nm is used to make an estimation of the helical content.[173] Table 3.4 presents the results of the helical content either without or with $(CF_3)_2CHOH$ for peptidomimetics **1**, **3**-**6** at 25 ˚C. Surprisingly, these data show that peptidomimetic **6** was the most helical of the set. This result suggests that the helicities of the peptide

sequences **1**, **3**-**5** are mainly from the contribution of the amino acids similar to those of peptidomimetic **6**.

Table 3.4. Summary of CD Data and Estimated Helical Contents for Peptidomimetics 1, 3-6

compound	20 mM phosphate buffer, pH 4		20 mM phosphate buffer with indicated vol % $(CF_3)_2CHOH$, pH 4		
□	$[□]_{222}$ (deg cm^2 $dmol^{-1}$)	helix (%)	$(CF_3)_2CHOH$ (%)	$[□]_{222}$ (deg cm^2 $dmol^{-1}$)	helix (%)
1	-8086	28	20	-12199	43
3	-8601	30	15	-14844	52
4	-8385	27	15	-16588	53
5	-9082	29	15	-19638	62
6	-10679	62	15	-16210	94

(a)



(b)

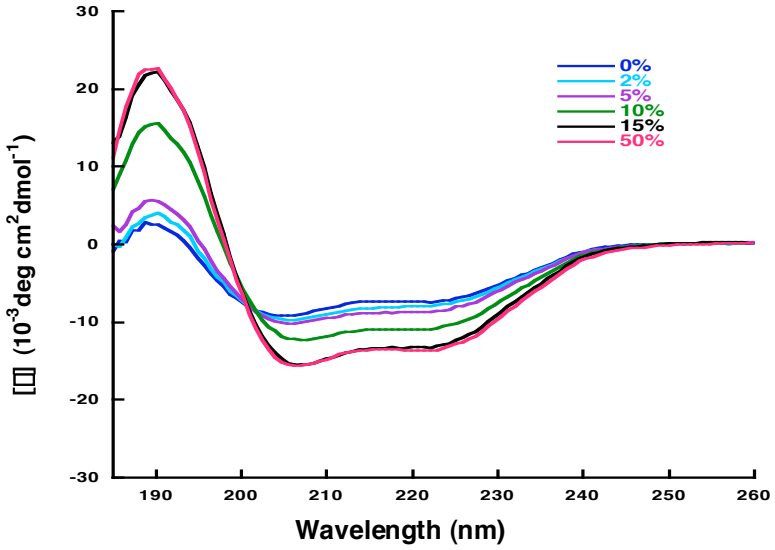
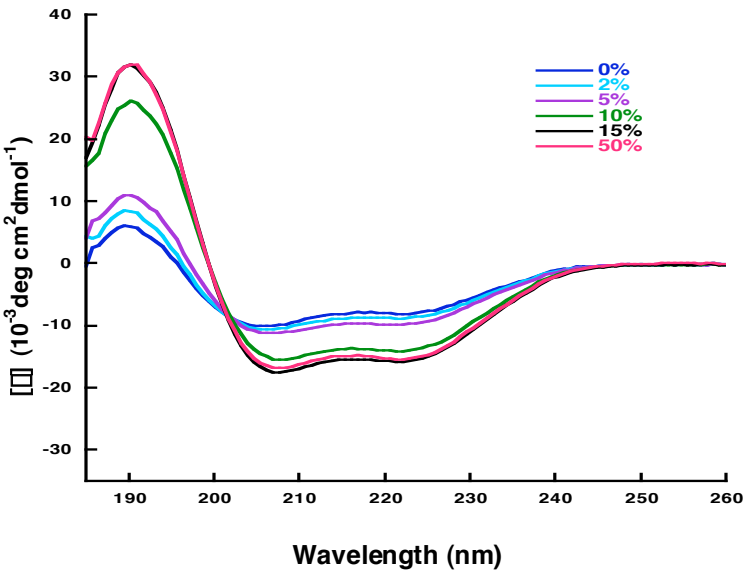


Figure 3.9. (a) - (b) CD spectra for compounds **1** and **3,** respectively, at 50 μM concentrations in pH 4 phosphate buffer (20 mM), and in the indicated buffer:$(CF_3)_2CHOH$ ratios (by volume).

(a)



(b)

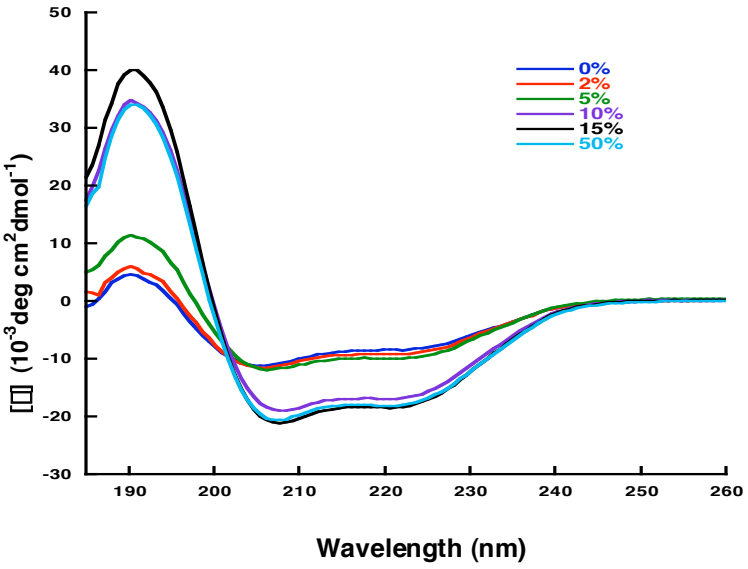
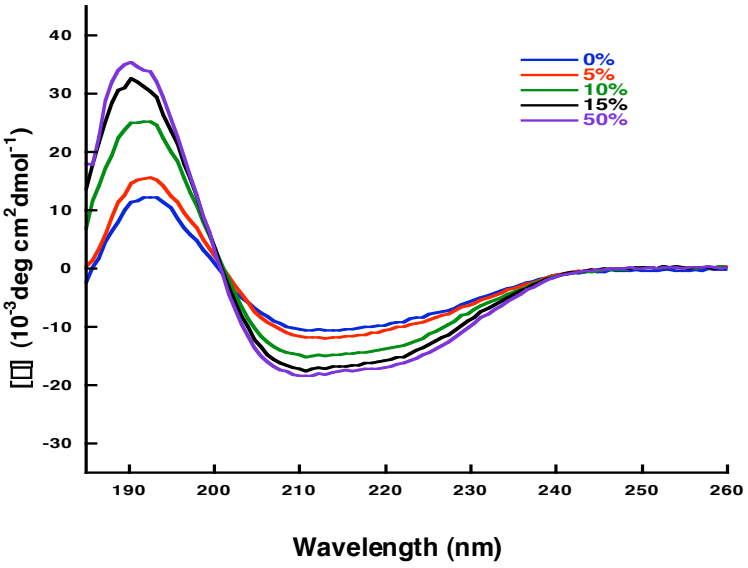


Figure 3.10. (a) - (b) CD spectra for compounds **4** and **5,** respectively, at 50 μM concentrations in pH 4 phosphate buffer (20 mM), and in the indicated buffer:$(CF_3)_2CHOH$ ratios (by volume).

(a)



(b)

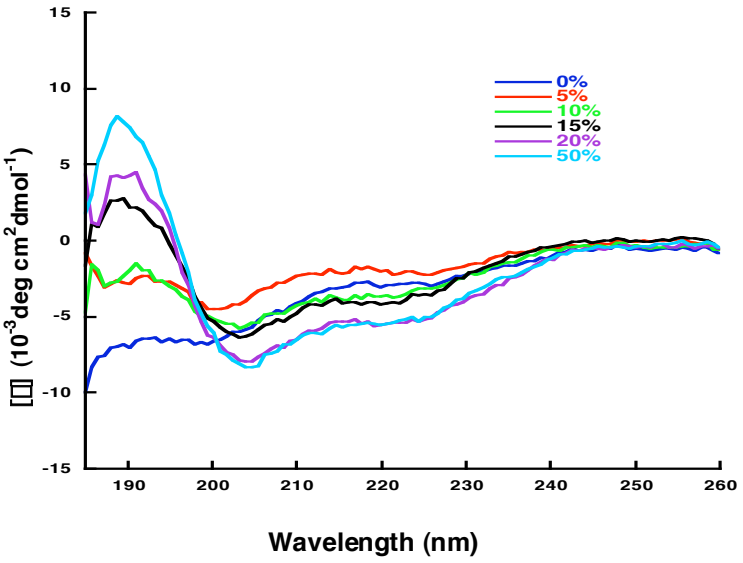


Figure 3.11. (a) - (b) CD spectra for compounds **6** and **7,** respectively, at 50 μM concentrations in pH 4 phosphate buffer (20 mM), and in the indicated buffer:$(CF_3)_2CHOH$ ratios (by volume).

In addition to the CD measurement for each compound, the CD spectra of each compound mixed with IgG have also been recorded. This study is to know whether addition of IgG to the peptidomimetics improves or changes their □-helical secondary structure or not. Consequently, for peptidomimetics **1**-**5**, and **7**, there is no significant difference in CD spectra between peptidomimetics alone and the mixture of peptidomimetics with IgG. For peptidomimetic **6**, however, the addition of IgG increases the □-helical secondary structure content, from –10.679 to –12.324, corresponding to 62% and 72% (Figure 3.12).

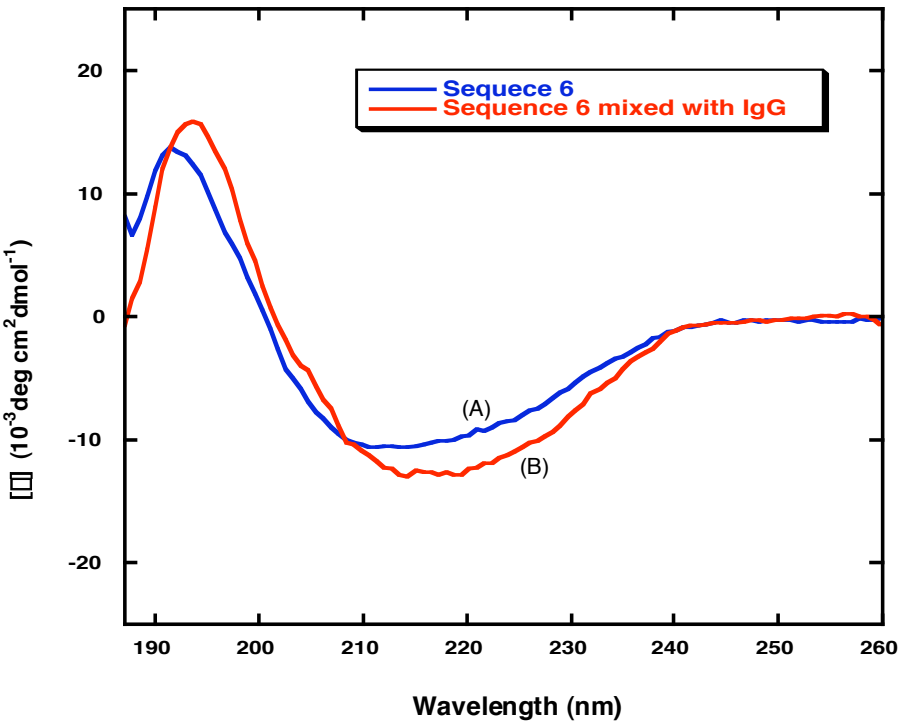


Figure 3.12. CD spectra for peptidomimetic **6** (50 □M) alone (A) and peptidomimetic **6** (6 □M) mixed with IgG (0.125 mg/ml) (B) in pH 4 phosphate buffer (20 mM).

3.5.2 NMR Analyses of Peptidomimetics

Three types of solvent systems were used in NMR studies. Specifically, DMSO-d6, 90% H_2O : 10% D_2O , and 80% H_2O : 20% $(CF_3)_2CDOD$ were used for compound **1**; DMSO-d6, 90% H_2O : 10% D_2O , and 85% H_2O : 15% $(CF_3)_2CDOD$ were used for compounds **4**, **5** and **6**; DMSO-d6 was used for compound **2**, and 90% H_2O : 10% D_2O was used for compound **3**; DMSO-d6 and 90% H_2O : 10% D_2O were used for compound **7**. The ratio of H_2O and $(CF_3)_2CDOD$ was determined by a series of CD experiments (Table 3.4). Whenever water was used, it was buffered with 20 mM sodium phosphate (pH = 4). Since compounds **1**, **2** and **3** are structurally similar, and compound **7** does not have any helical propensity based on CD data, compounds **1**, **4**-**6** were finally selected for NMR conformational studies.

3.5.2.1 Characterization of Seven Peptidomimetics

Chemical shifts of seven peptides were assigned by analyzing the spectra of 1D and 2D NMR including DQF-COSY,[211] TOCSY,[212] NOESY[151-153] or ROESY.[154,155] The spin systems of amino acids were identified by the through bond-coupling network of DQF-COSY and TOCSY spectra obtained. One cross peak for every αH-NH coupling of each amino acid residue is expected to appear in the fingerprint region of DQF-COSY and TOCSY spectra. The sequential assignments for peptides were carried out by the through-space NOE connectivities of NOESY or ROE connectivities of ROESY experiments. The amide region and fingerprint of the NOESY and ROESY spectra provide important information to verify the backbone structure of a peptide.

The solvent system (90% H_2O/10% D_2O) was chosen to illustrate the characterization of peptidomimetics **1**, **4**, **5** and **6**. The fingerprint region of TOCSY spectra, and the N<u>H</u>-N<u>H</u> and Cα/N<u>H</u> regions of ROESY or NOESY spectra of these compounds were shown in Figures 3.13-3.20. The chemical shifts and coupling constants of peptidomimetics **1**, **4-6** were tabulated in Tables 3.5-3.8 respectively.

mimic **1**

Succinyl-Asp-Phe-Tyr-Aib-Ile-Leu-His-Acp-Gln-His-Asn-Aib-Phe-Ile-Asn-NH_2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

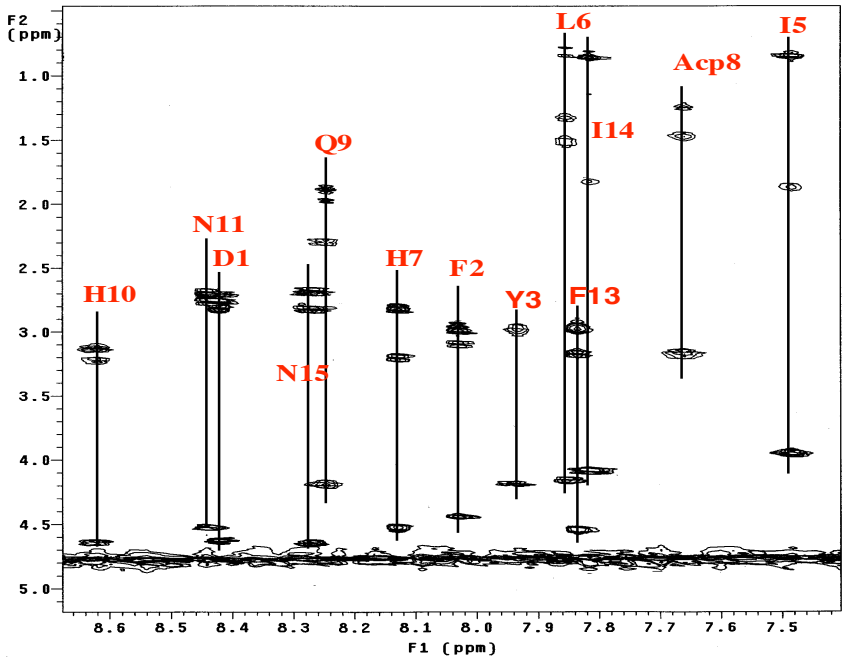
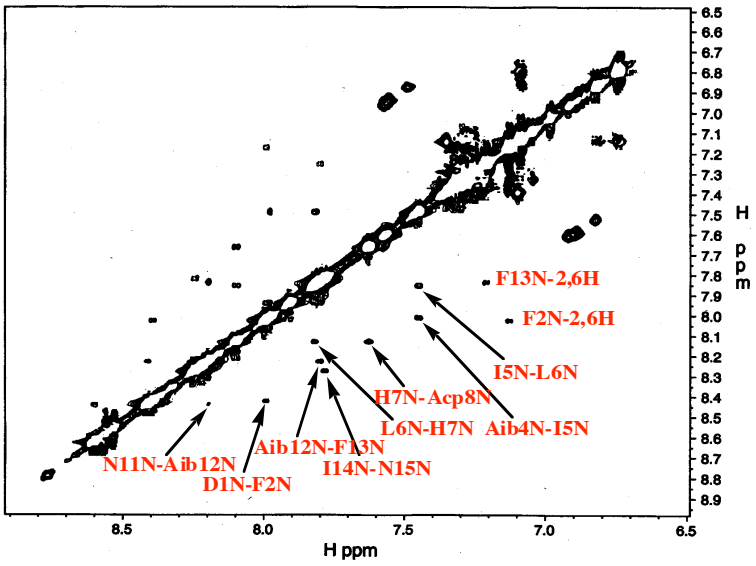


Figure 3.13. Fingerprint region of a 80-ms TOCSY spectrum of mimic **1** in 90% H_2O/10% D_2O. Individual spin systems are identified with one-letter symbols for the amino acid residues.

(a)



(b)

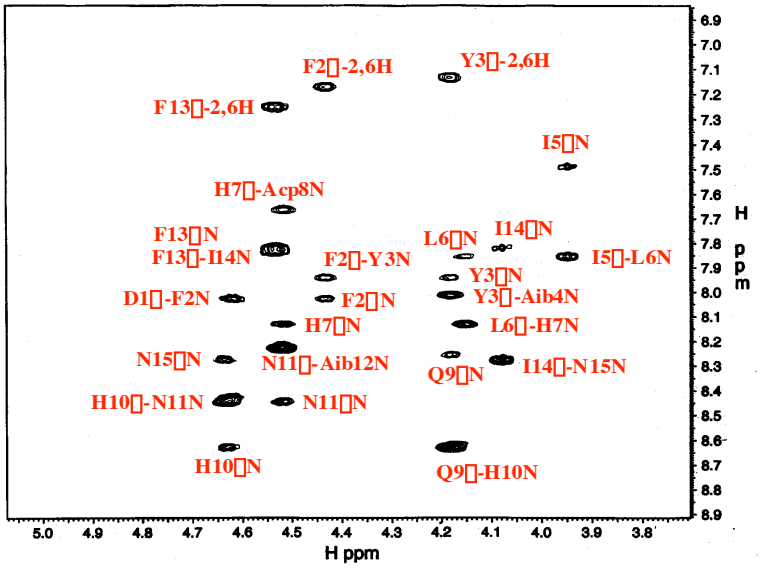


Figure 3.14. (a) - (b) NH-NH and C/NH "fingerprint" regions, respectively, of a 150-ms ROESY spectrum of mimic **1** in 90% H_2O/10% D_2O, pH=4, 25 °C.

Table 3.5. Chemical Shifts and Coupling Constants of Mimic 1 (in 90% H_2O/10% D_2O)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.41/6.2	4.63	2.82, 2.74			
Phe-2	8.02/5.6	4.44	3.10, 2.97			3,4,5H:7.34-7.25(m), 2,6H:7.13(d,J=6.5)
Tyr-3	7.93/4.8	4.18	2.98, 2.98			2,6H: 7.09(d,J=7.5), 3,5H: 6.74(d,J=7.5)
Aib-4	7.83		1.32, 1.24			
Ile-5	7.48/6.0	3.96	1.88	1.46, 1.21, 0.85	0.82	
Leu-6	7.85/5.5	4.17	1.52, 1.52	1.33	0.84, 0.78	
His-7	8.12/6.9	4.53	3.21, 2.81			2H: 8.54, 4H: 7.02
Acp-8	7.65/5.5					$(CH_2)_5$: 3.17, 3.17, 2.24, 1.47, 1.24
Gln-9	8.24/5.75	4.2	1.97, 1.88	2.30, 2.30		εNH_2: 7.51, 6.83
His-10	8.62/7.0	4.64	3.22, 3.13			2H: 8.57, 4H: 7.28
Asn-11	8.43/5.9	4.52	2.78, 2.69			δNH_2: 7.58, 6.90
Aib-12	8.21		1.31, 1.23			
Phe-13	7.82/7.65	4.54	3.17, 2.97			3,4,5H: 7.34-7.25(m), 2,6H: 7.21(d,J=6.5)
Ile-14	7.81/7.5	4.09	1.82	1.43, 1.13, 0.86	0.82	
Asn-15	8.26/7.0	4.65	2.82, 2.68			δNH_2: 7.60, 6.93
Succinyl						CH_2: 2.63, CH_2: 2.51
NH_2	7.38, 7.10					

mimic **4**

Succinyl-Asp-Gln-Gln-Asn-Ala-Phe-Tyr-Aib-Ala-Leu-His-Acp-Ala-Arg-Asn-Ala-Aib-Ile-Gln-Asn-NH_2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

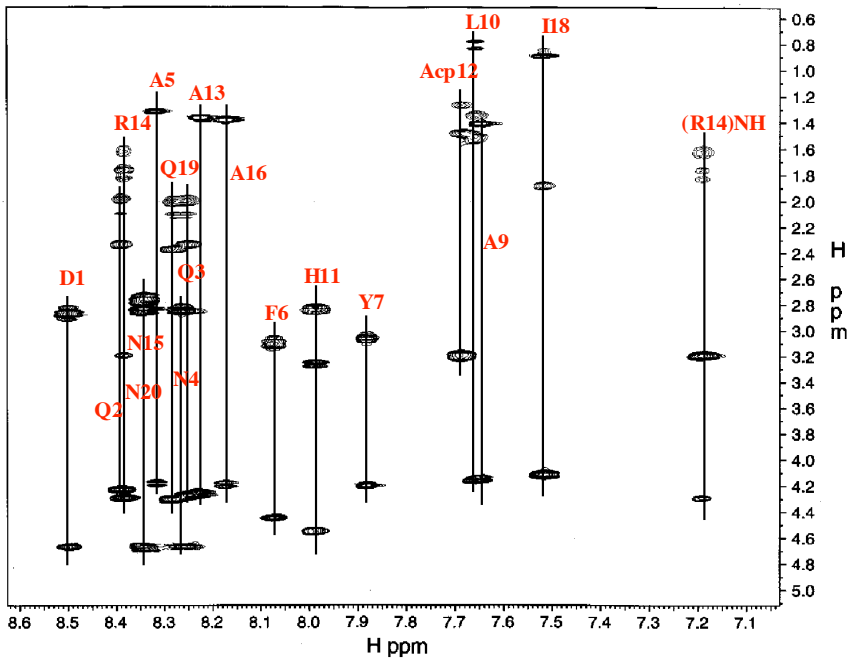
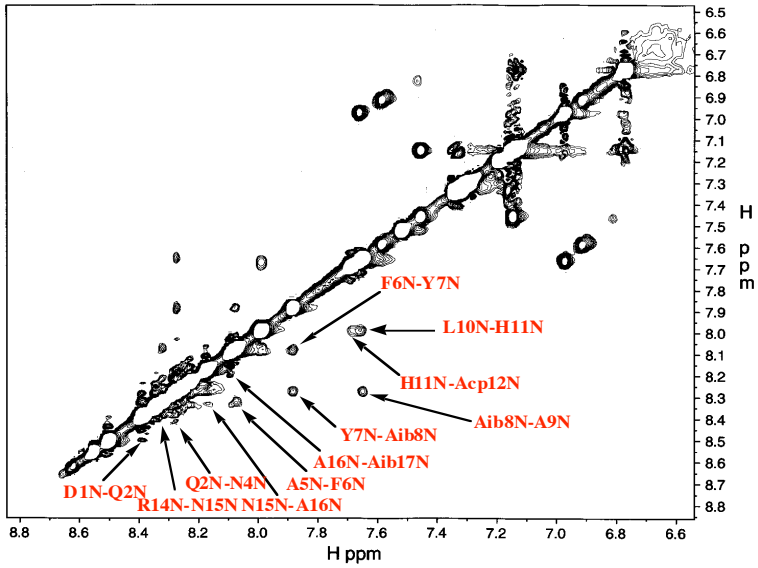


Figure 3.15. Fingerprint region of a 80-ms TOCSY spectrum of mimic **4** in 90% H_2O/10% D_2O. Individual spin systems are identified with one-letter symbols for the amino acid residues.

(a)



(b)

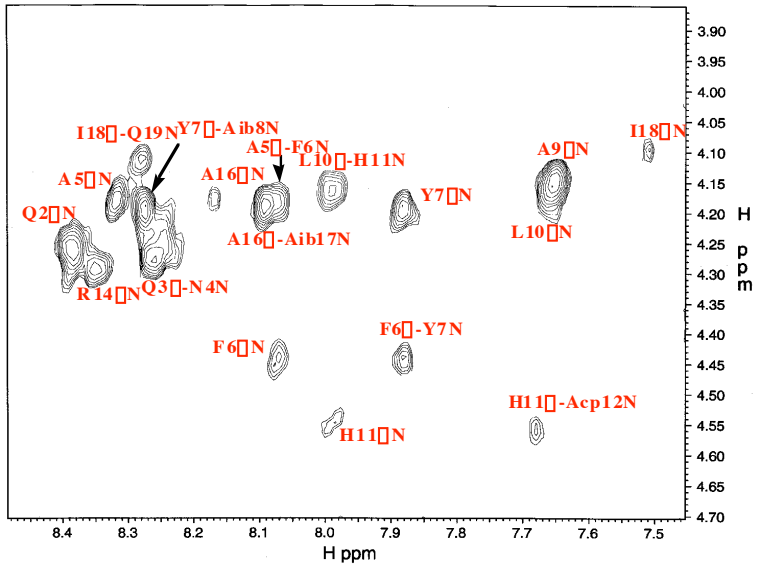


Figure 3.16. (a) - (b) NH-NH and C/NH "fingerprint" regions, respectively, of a 300-ms NOESY spectrum of mimic **4** in 90% H_2O/10% D_2O, pH=4, 25 °C.

Table 3.6. Chemical Shifts and Coupling Constants of Mimic 4 (in 90% H_2O/10% D_2O)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.54/6.2	4.68	2.88, 2.88			
Gln-2	8.43/5.8	4.24	2.15, 2.01	2.37, 2.37		δNH_2: 7.69, 7.01
Gln-3	8.29/5.0	4.27	2.11, 1.98	2.35, 2.35		δNH_2: 7.54, 6.88
Asn-4	8.30/6.5	4.68	2.87, 2.87			δNH_2: 7.50, 6.85
Ala-5	8.35/4.45	4.18	1.33			
Phe-6	8.11/5.65	4.47	3.13, 3.13			3,4,5H:7.31-7.39(m), 2,6H:7.19 (d,J=7.0)
Tyr-7	7.92/5.1	4.21	3.09, 3.09			2,6H: 7.17(d,J=7.5), 3,5H: 6.81(d,J=7.5)
Aib-8	8.31		1.38, 1.31			
Ala-9	7.68/4.45	4.13	1.44			
Leu-10	7.70/5.5	4.18	1.59, 1.59	1.36	0.84, 0.80	
His-11	8.02/8.0	4.57	3.29, 2.85			2H: 8.59, 4H: 7.03
Acp-12	7.73/5.5					$(CH_2)_5$: 3.22, 2.27, 1.57, 1.48, 1.28
Ala-13	8.26/4.95	4.25	1.39			
Arg-14	8.42/6.0	4.31	1.82, 1.76	1.61, 1.61	3.18, 3.18	NH: 7.26 (t, J=5.5)
Asn-15	8.38/7.5	4.68	2.89, 2.75			δNH_2: 7.62, 6.95
Ala-16	8.21/4.95	4.18	1.41			
Aib-17	8.13		1.48, 1.39			
Ile-18	7.55/7.0	4.11	1.91	1.42, 1.19, 0.91	0.91	
Gln-19	8.32/6.5	4.33	2.11, 1.99	2.41, 2.41		δNH_2: 7.54, 6.88
Asn-20	8.38/7.0	4.68	2.89, 2.75			δNH_2: 7.61, 6.94
Succinyl						CH_2 2.56, CH_2 2.47
NH_2	7.50, 7.18					

mimic **5**

Succinyl-Asp-Ala-Phe-Tyr-Aib-Ala-Leu-His-Acp-Ala-Arg-Asn-Ala-Aib-Ile-Gln-Ser-Leu-Lys-Asn-NH_2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

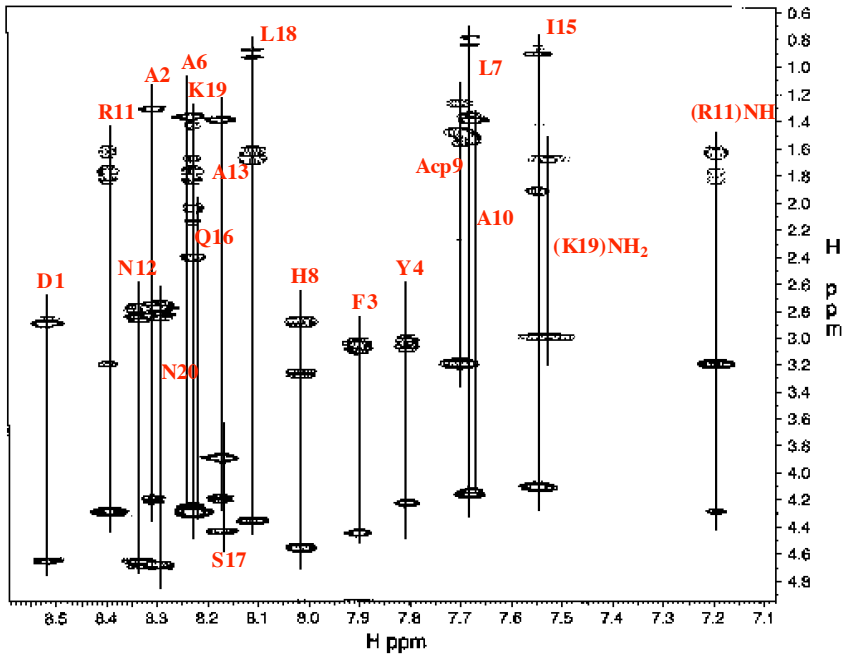


Figure 3.17. Fingerprint region of a 80-ms TOCSY spectrum of mimic **5** in 90% H_2O/10% D_2O. Individual spin systems are identified with one-letter symbols for the amino acid residues.

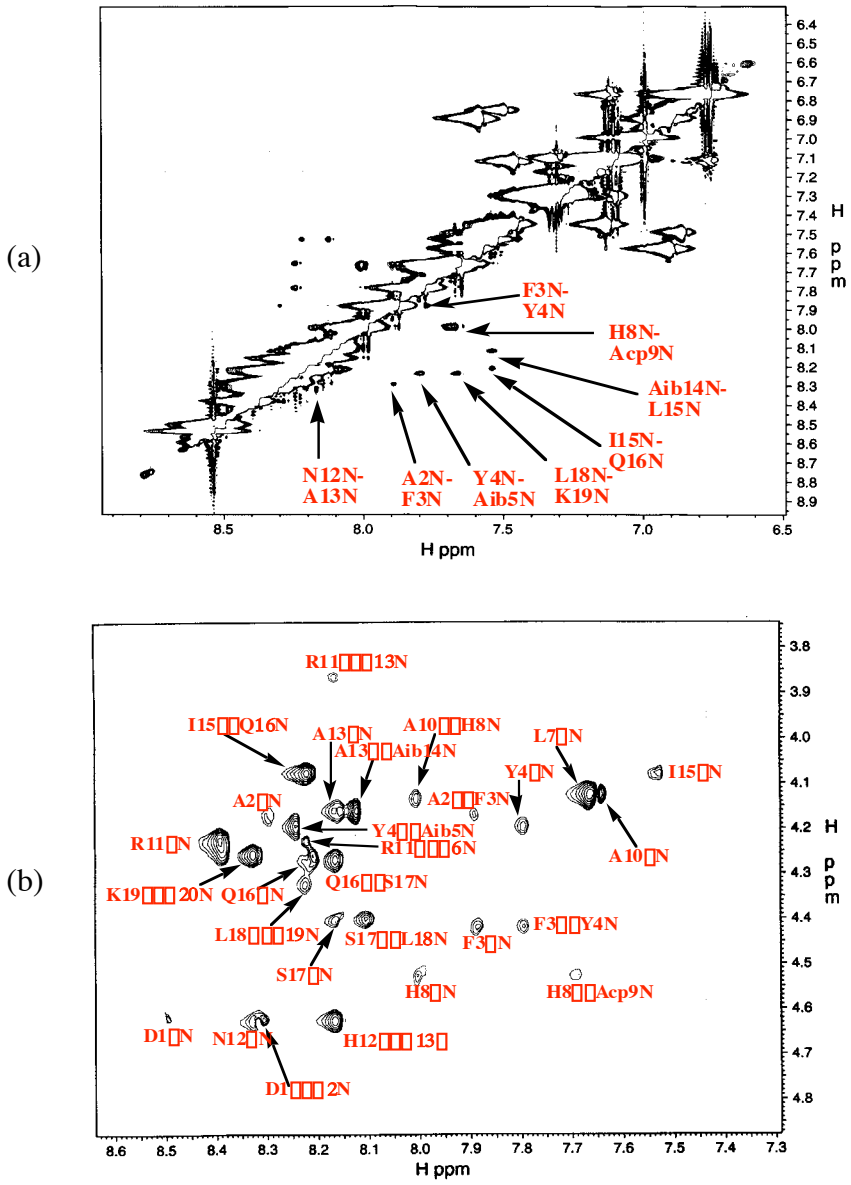


Figure 3.18. (a) - (b) NH-NH and C□/NH "fingerprint" regions, respectively, of a 300-ms NOESY spectrum of mimic **5** in 90% H_2O/10% D_2O, pH=4, 25 ºC.

Table 3.7. Chemical Shifts and Coupling Constants of Mimic 5 (in 90% H_2O/10% D_2O)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.50/5.8	4.52	2.88, 2.88			
Ala-2	8.28/4.6	4.17	1.3			
Phe-3	7.87/5.9	4.43	3.07, 3.01			3,4,5H:7.32-7.25(m), 2,6H:7.09 (d,J=8.5)
Tyr-4	7.78/5.5	4.2	3.04, 3.01			2,6H: 7.11(d,J=8.5), 3,5H: 6.76(d,J=8.5)
Aib-5	8.23		1.43, 1.38			
Ala-6	8.21/4.65	4.18	1.39			
Leu-7	7.65/6.5	4.14	1.52, 1.52	1.34	0.82, 0.76	
His-8	7.99/8.0	4.54	3.25, 2.86			2H: 8.53, 4H: 6.99
Acp-9	8.67/5.5					$(CH_2)_5$: 3.18, 3.18, 2.24, 1.52, 1.25
Ala-10	7.65/5.0	4.13	1.35			
Arg-11	8.37/5.9	4.27	1.83, 1.76	1.61, 1.61	3.18, 3.18	NH: 7.17 (t, J=5.5)
Asn-12	8.31/7.0	4.63	2.84, 2.76			δNH_2: 7.57, 6.89
Ala-13	8.15/4.9	4.15	1.4			
Aib-14	8.11		1.44, 1.36			
Ile-15	7.52/6.5	4.08	1.9	1.44, 1.19, 0.90	0.85	
Gln-16	8.20/4.15	4.26	2.12, 2.03	2.38, 2.38		δNH_2: 7.48, 6.85
Ser-17	8.15/7.0	4.4	3.87, 3.87			
Leu-18	8.09/6.5	4.34	1.67, 1.67	1.6	0.92, 0.86	
Lys-19	8.20/	4.26	1.82, 1.76	1.41, 1.41	1.66, 1.66	δCH_2: 2.76, 2.76, NH_2: 7.49, 7.10
Asn-20	8.27/7.5	4.65	2.82, 2.75			δNH_2: 7.57, 6.89
Succinyl						CH_2 2.40, CH_2 2.25
NH_2	7.44, 7.12					

mimic **6**

Succinyl-Asp-Phe-Tyr-Aib-Ile-Leu-His-NH_2

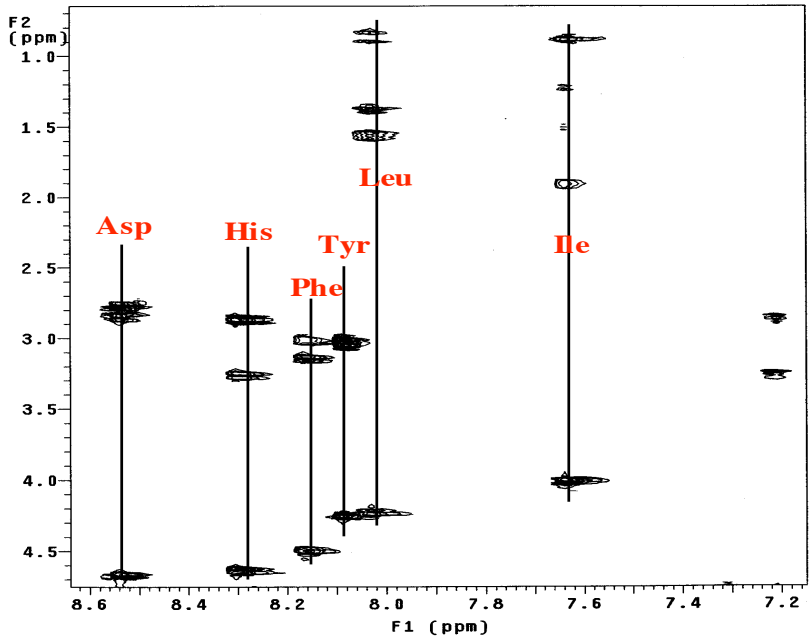


Figure 3.19. Fingerprint region of a 80-ms TOCSY spectrum of mimic **6** in 90% H_2O/10% D_2O. Individual spin systems are identified with three-letter symbols for the amino acid residues.

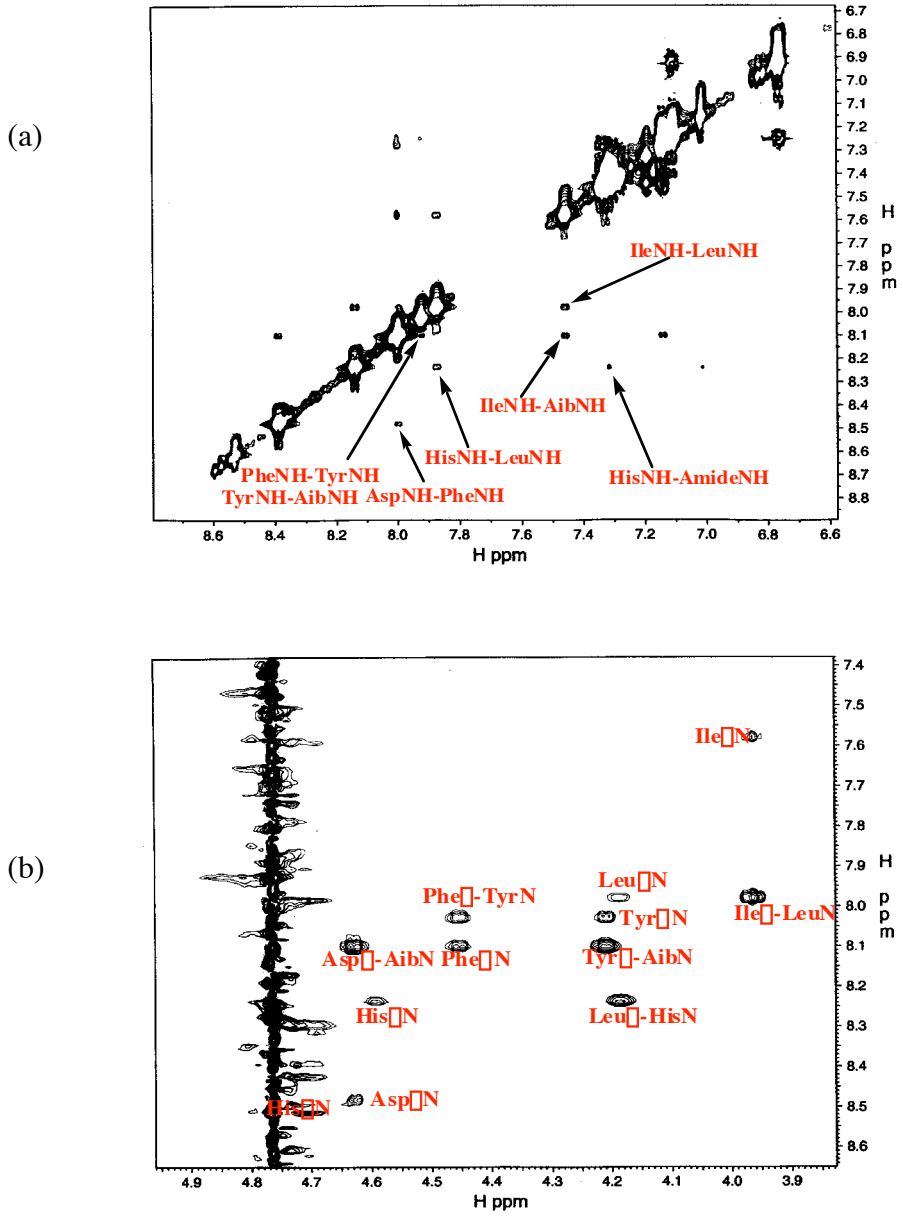


Figure 3.20. (a) - (b) NH-NH and C/NH "fingerprint" regions, respectively, of a 150-ms ROESY spectrum of mimic **6** in 90% H_2O/10% D_2O, pH=4, 25 °C.

Table 3.8. Chemical Shifts and Coupling Constants of Mimic 6 (in 90% H_2O/10% D_2O)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp	8.51/5.9	4.68	2.86, 2.77			
Phe	8.15/4.7	4.5	3.16, 3.02			3,4,5H:7.52-7.42(m), 2,6H:7.30(d,J=7.5)
Tyr	8.08/4.8	4.37	3.06, 3.01			2,6H: 7.26(d,J=8.5), 3,5H: 6.92(d,J=8.5)
Aib	8.15		1.45, 1.46			
Ile	7.61/5.5	4.12	1.91	1.50, 1.25, 0.92	0.88	
Leu	8.03/5.5	4.24	1.58, 1.58	1.4	0.90, 0.85	
His	8.26/7.0	4.65	3.27, 2.78			2H: 8.68, 4H: 7.17
Succinyl						CH_2: 2.70, CH_2: 2.57
NH_2	7.48, 7.36					

3.5.2.2 □-Helical Conformation Observed via NMR Spectroscopies in Solution

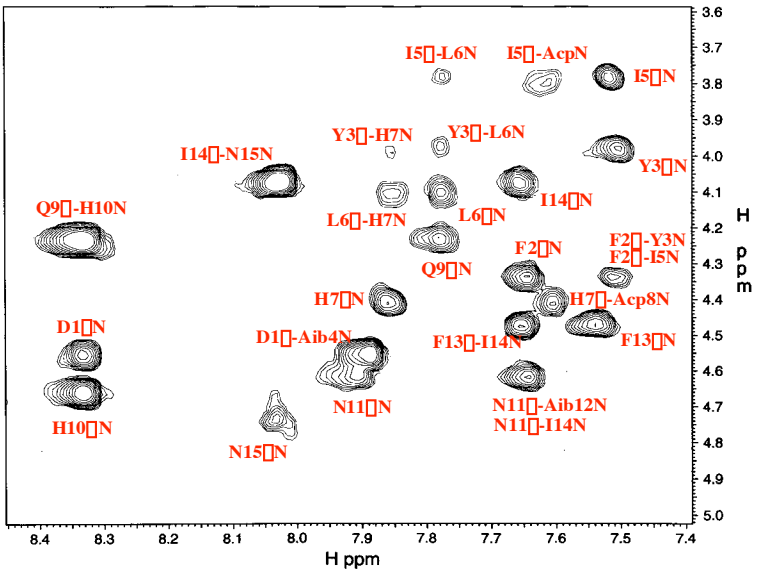
The NMR studies were performed in 1:9 D_2O : buffer medium (20 mM phosphate), and in $(CF_3)_2CDOD$: buffer using the same ratios used for the CD studies indicated in Table 3.4. The following discussion only relates to the $(CF_3)_2CDOD$: buffer media since the conformational preferences were more pronounced in that solvent system. Chemical shifts (see Appendix E) were assigned via a combination of DQF-COSY, TOCSY, and NOESY (**4** and **5**) or ROESY (**1** and **6**) spectroscopies. The latter techniques were then used to detect close contacts of protons. N<u>H</u>-N<u>H</u> and C□/N<u>H</u> "fingerprint" regions of the NOESY or ROESY spectra of compounds **1**, **4** □ **6** were shown in Figures 3.21-3.24. Intraresidue: $d_{\alpha N}(i,i)$, sequential: $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$, and medium-range: $d_{\alpha N}(i,i+3)$ connectivities are denoted. The close contacts of protons, along with coupling constant information, are summarized in Figure 3.25. Figure 3.7 shows the crosspeaks that should be observed for ideal helical conformations. Comparison of the data presented in Figure 3.25 with this shows that the *N*-terminal sequence for each of the model peptides has a more distinct conformational bias to helical character than the *C*-terminal helix. Even peptidomimetic **6**, that includes the residues for the *C*-terminal helical part, shows a pronounced tendency towards helical character in the other {*N*-terminal} region. No long-range crosspeaks were observed, again consistent with helical structures. These assertions are supported by data from coupling constants that show several small $^3J_{H\alpha,HN}$ values (< 6 Hz) (see Appendix E).

mimic **1**

Succinyl-Asp-Phe-Tyr-Aib-Ile-Leu-His-Acp-Gln-His-Asn-Aib-Phe-Ile-Asn-NH_2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

(a)



(b)

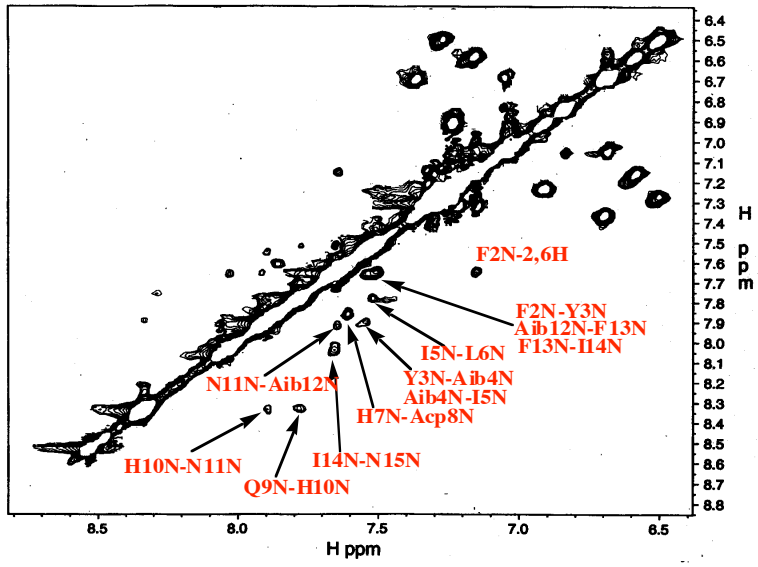


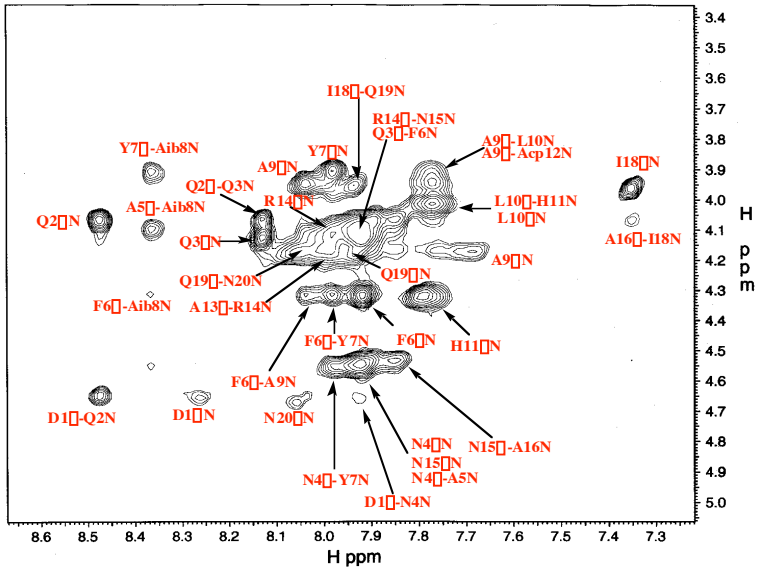
Figure 3.21. (a) - (b) C□/N<u>H</u> and N<u>H</u>-N<u>H</u> "fingerprint" regions, respectively, of a 200-ms ROESY spectrum of mimic **1** in 80% buffer/20% $(CF_3)_2CDOD$, pH=4, 25 °C.

mimic **4**

Succinyl-Asp-Gln-Gln-Asn-Ala-Phe-Tyr-Aib-Ala-Leu-His-Acp-Ala-Arg-Asn-Ala-Aib-Ile-Gln-Asn-NH_2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

(a)



(b)

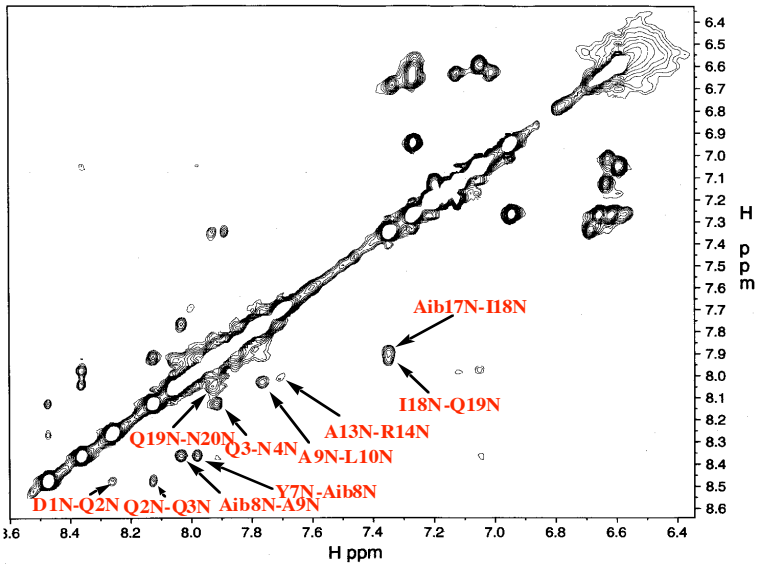


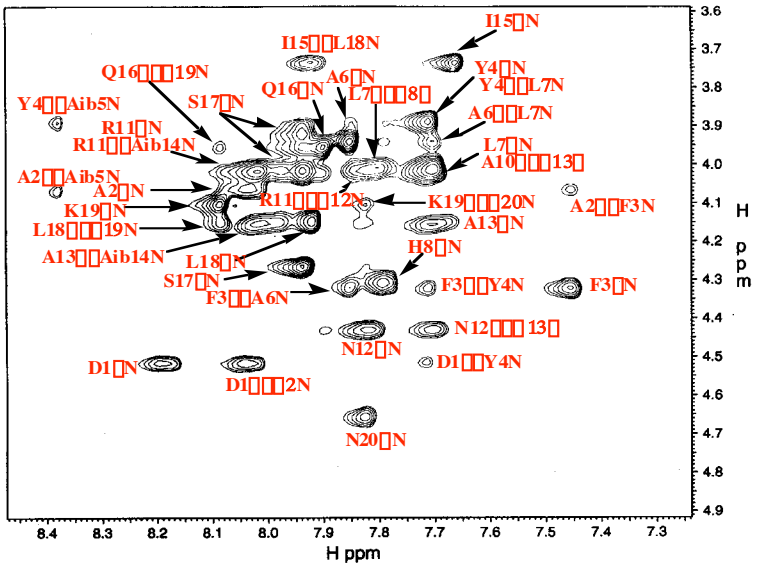
Figure 3.22. (a) - (b) C□/N<u>H</u> and N<u>H</u>-N<u>H</u> "fingerprint" regions, respectively, of a 250-ms ROESY spectrum of mimic **4** in 85% buffer/15% $(CF_3)_2CDOD$, pH=4, 25 °C.

mimic **5**

Succinyl-Asp-Ala-Phe-Tyr-Aib-Ala-Leu-His-Acp-Ala-Arg-Asn-Ala-Aib-Ile-Gln-Ser-Leu-Lys-Asn-NH_2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

(a)



(b)

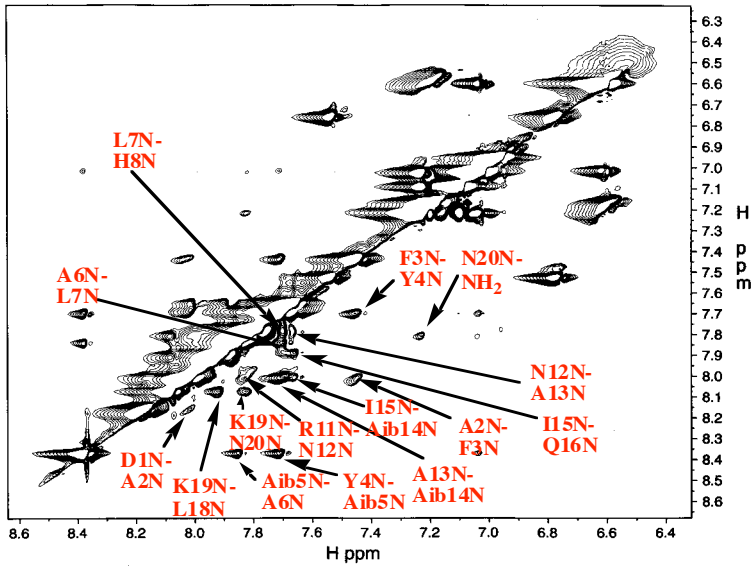
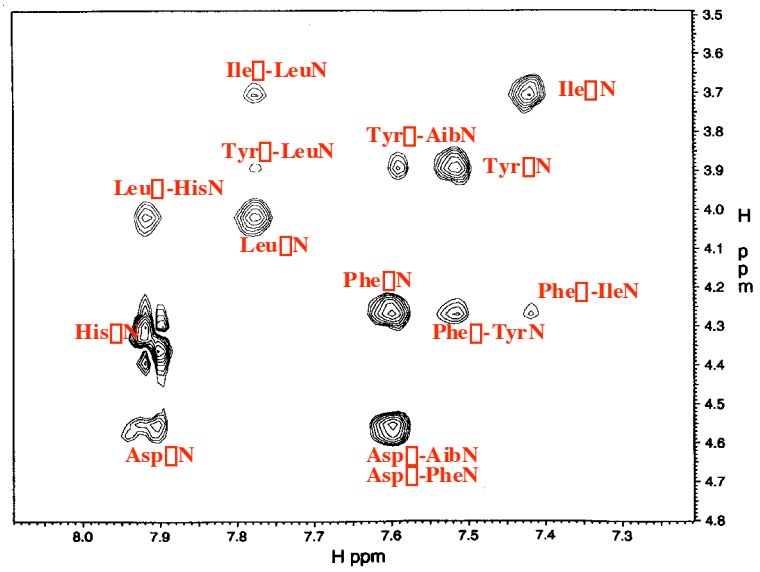


Figure 3.23. (a) - (b) C□/N<u>H</u> and N<u>H</u>-N<u>H</u> "fingerprint" regions, respectively, of a 250-ms ROESY spectrum of mimic **5** in 85% buffer/15% $(CF_3)_2CDOD$, pH=4, 25 ºC.

mimic **6**

Succinyl-Asp-Phe-Tyr-Aib-Ile-Leu-His-NH_2

(a)



(b)

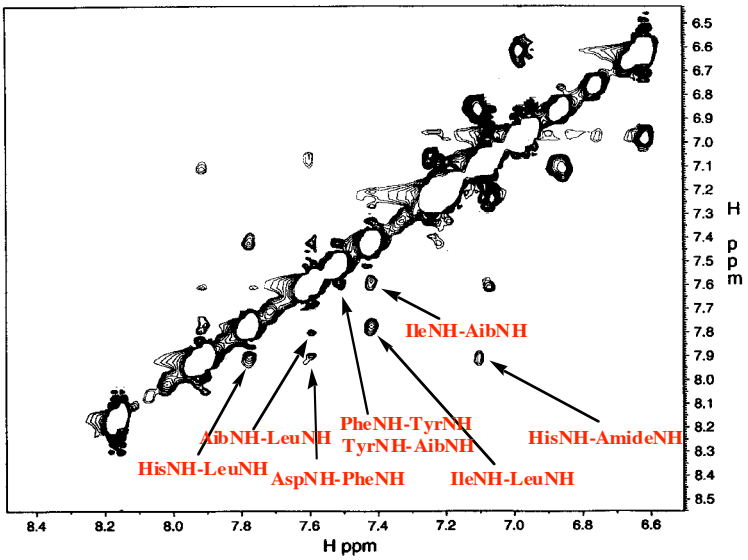
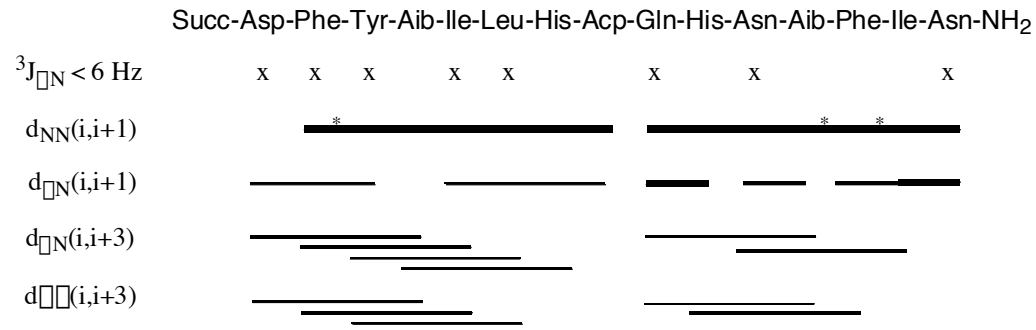


Figure 3.24. (a) - (b) CαH/NH and NH-NH "fingerprint" regions, respectively, of a 200-ms ROESY spectrum of mimic **6** in 85% buffer/15% $(CF_3)_2CDOD$, pH=4, 25 °C.

(a)



(b)

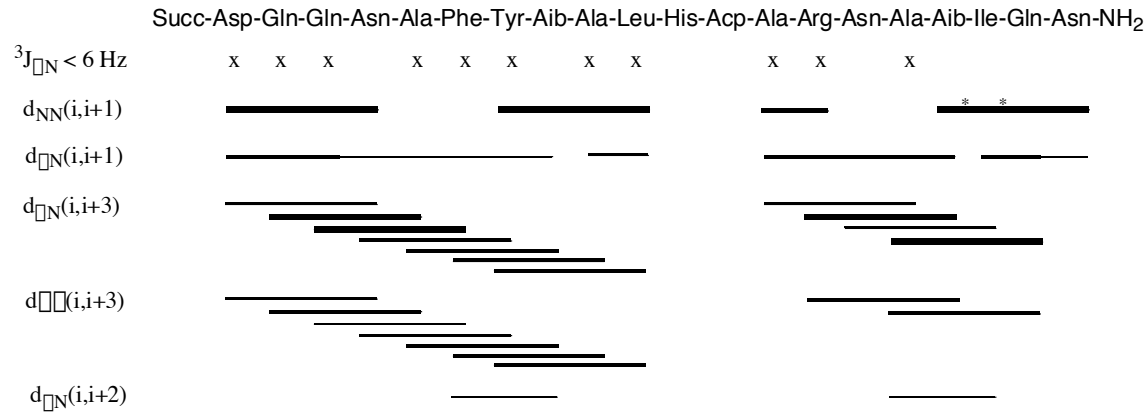
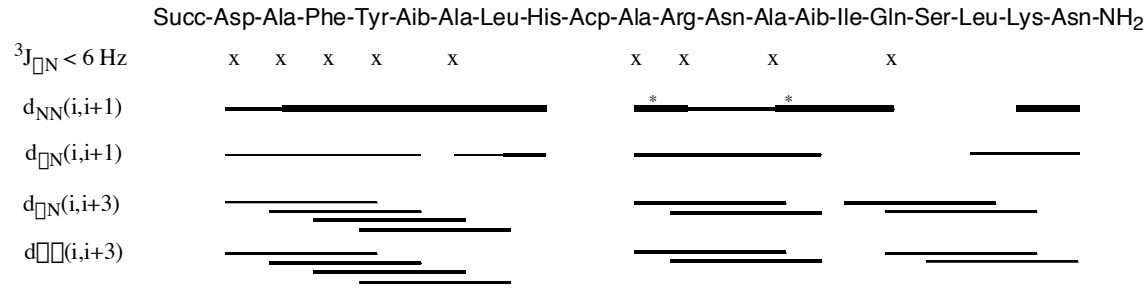


Figure 3.25. (a) - (d) Summary of the sequential and medium-range connectivities for peptidomimetics **1**, **4**-**6** respectively. The intensities of NOE/ROE cross-peaks are indicated by the thickness of the lines. Where two resonances overlap so absolute assignments were not possible, this is indicated with an asterisk (*).

(c)



(d)

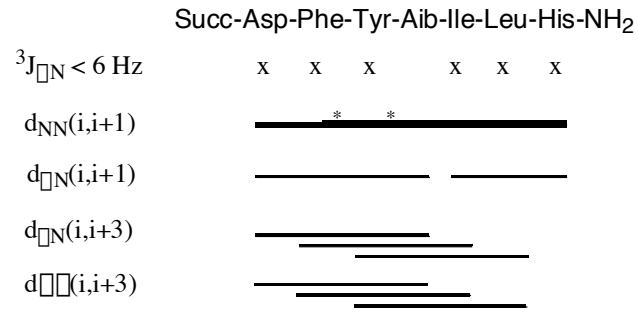


Figure 3.25. (Continued).

When 1:9 D_2O:buffer was used as the NMR solvent system, several short and medium-range ROE or NOE connectivities were observed (Figures 3.14, 3.16, 3.18 and 3.20). The ROE or NOE connectivities and a number of small $^3J_{H\alpha,HN}$ coupling constants (< 6 Hz) (Tables 3.5-3.8) indicated a partial β-helical conformation within peptidomimetics **1**, **4** – **6**. The β-helix propensity is mainly contributed from the fragment of peptidomimetics **1**, **4** – **6**, which are corresponding to the first helical region. After $(CF_3)_2CDOD$ substitutes for D_2O, based on the ratios (Table 3.4) for NMR solvent system, the number of short and medium-range ROE or NOE connectivities increases (Figures 3.21-3.24). Thus addition of hexafluoroisopropanol to the medium results in a stabilization of an α-helix of peptidomimetics **1**, **4** – **6**. No long-range ROE connectivities were observed either before or after addition of hexafluoroisopropanol.

3.5.2.3 Secondary Structures of Peptidomimetics 1, 4-6 Generated from Constrained Molecular Dynamics

The ROE/NOE connectivity data and dihedral angles (ϕ) deduced from coupling constants were used as distance and torsion angle restraints in constrained molecular dynamics studies. There were performed using the Quanta/CHARMm package following a protocol similar to that described for quenched molecular dynamics (QMD)[79,213] studies. Thus the NMR data were obtained to generate sets of molecular coordinates, each one corresponding to an instantaneous conformational state of the molecules studied. Overall these data give an impression of the most populated conformational states that are consistent with the restraints applied from the NMR data. Figure 3.26 presents the lowest energy structures of peptidomimetics **1**, **4**, **5**, and **6** that were identified within 6 - 7 kcal/mol of the low energy conformation for that particular peptidomimetic. Based on this simulation, all the peptidomimetics may populate helical states at the *N*-terminus, but the *C*-terminal conformations of the residues designed to mimic the second helical region in **1**, **4**, and **5** display little or not helical character.

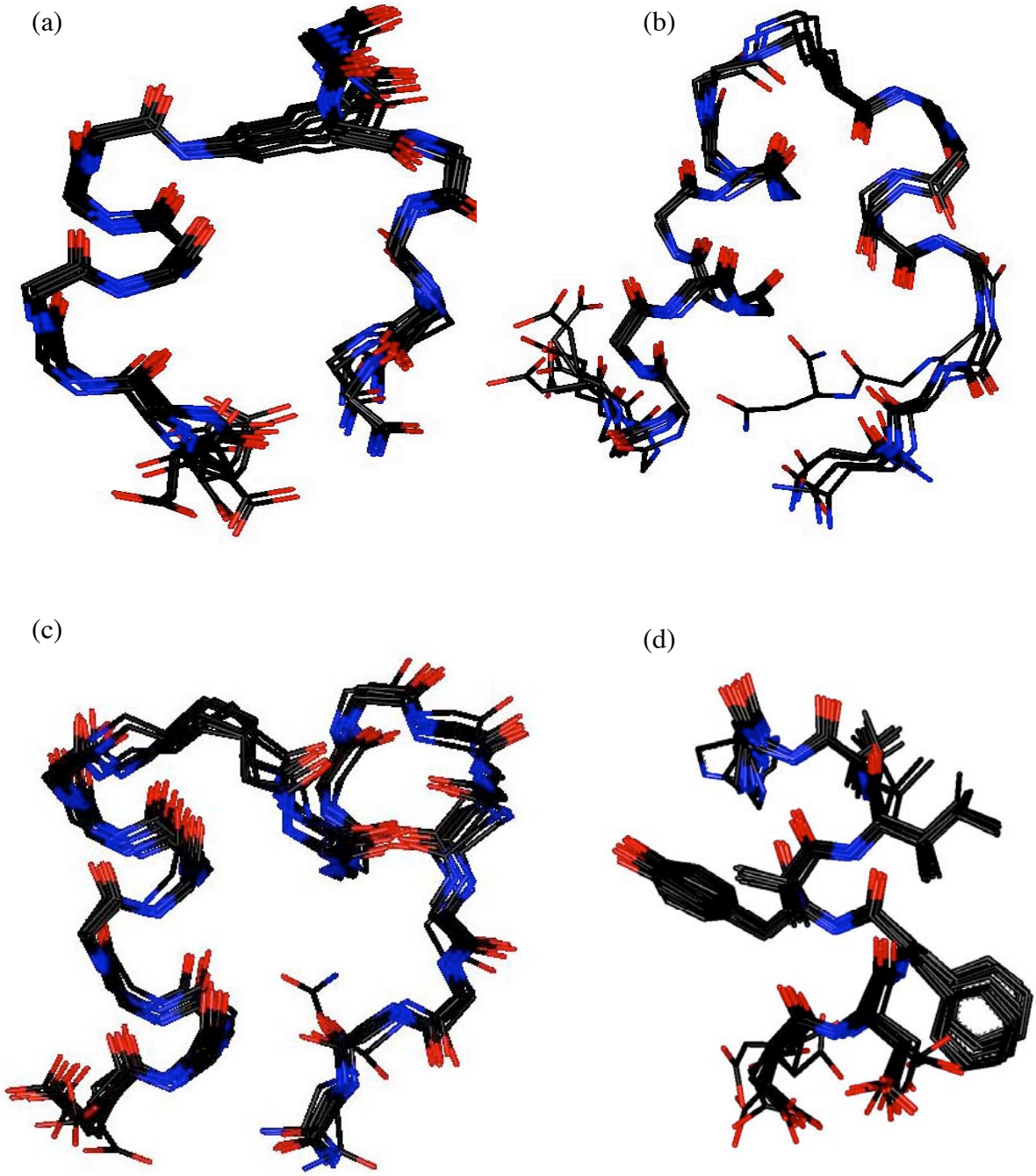


Figure 3.26. (a) - (d) Superimposed low energy structures generated for peptidomimetics **1** (11), **4** (6), **5** (10), and **6** (12) {numbers of structures within the cut-off threshold are indicated in parentheses}, respectively. The side chains were omitted for clarity in peptidomimetics **1**, **4** and **5**.

3.6 Summary

The structures of seven peptidomimetics have been characterized by NMR and MALDI-MS. The □-helical secondary structures of peptidomimetics **1**, **4**-**6** have been characterized by NMR spectroscopies and CD spectra. Consequently, peptidomimetics **1**, **4** and **5** have some helical character, which are mainly in the residues near the *N*-terminus. Peptidomimetic **6** with seven amino acids is the best one among seven mimics based on helical content. The NMR spectroscopies of peptidomimetics show sequential and some typical medium-range ROE or NOE connectivities, which signify an □-helical conformation. Small $^{3}J_{H\square, HN}$ coupling constants are additional supporting evidence.

$(CF_3)_2CHOH$, like trifluoroethanol, is a well-known helix-inducer. Addition of $(CF_3)_2CHOH$ to peptidomimetics stabilized their □-helical secondary structure, especially for peptidomimetic **6** whose helical content enhances from 62% to 94%, as observed in NMR and CD spectra. The evidence has shown that trifluoroethanol induces □-helical conformation for the peptides or peptide fragments which have an inherent propensity for such conformations.[171,214] In terms of this observation, peptidomimetic **6** probably has an inherent □-helix propensity. The residues close to the *N*-terminal of mimics **1**, **4** and **5** are structured better than those close to the *C*-terminal in the presence of $(CF_3)_2CHOH$. This indicates that the former has a higher □-helix propensity than the latter does.

CD studies provide an approach to see if a receptor improves the □-helix propensity of a ligand via their interaction. This is an interesting study. IgG has been proven to increase □-helical content of peptidomimetic **6** via this approach.

For the binding studies between ligands and a receptor, like IgG, via STD NMR spectroscopy, compounds **6** and **7** with IgG respectively were studies. There are no standard compounds, which shows good binding to IgG, for comparison, it is hard to make conclusions for my results. Mimics **1-7** were sent to Amersham Biosciences for bioactivity studies. Unfortunately, there were no binding activities to be found for these compounds.

CHAPTER IV
MONO- AND MULTIVALENT SMALL MOLECULE PEPTIDOMIMETICS OF PROTEIN G FOR AFFINITY SUPPORTS

4.1 Specific Aims

The goal of this project is to synthesize some monovalent small molecules designed by Amersham Pharmacia, and to design and synthesize our own multivalent small molecules. Both groups of small molecules are intended to mimic protein G that is also currently used to purify IgG. This research, specifically, attempts to:

a. synthesize several monovalent molecules, one of which has shown some binding to IgG by Amersham's preliminary data;
b. find a scaffold, and synthesize multivalent mimics based on Amersham's designed ligands and my peptidomimetic **6** discussed in Chapter III;
c. design our own monovalent compounds and synthesize building blocks;
d. develop a solid-phase methodology which can be used to attach building blocks to the scaffold efficiently;
e. make a small library of trivalent compounds via a parallel synthetic approach;
f. characterize all mimics by NMR, MS, etc.;
g. attach multivalent compounds onto sepharose for bioassay.

4.2 Background and Significance

4.2.1 Protein G and Its Binding Domain for IgG

Protein G is a small globular protein produced by groups C and G *Streptococci*. This globular protein consists of several nearly identical binding domains, *ie* III, B1, B2 and C2. The structures of these domains have been determined by means of X-ray crystallography[215-217] and NMR spectroscopy.[218,219] Crystal structures of the

protein G binding domain complexed with the Fc fragment (PDB: 1FCC)[220] and the Fab (PDB: 1IGC)[217] of IgG are known. Figures 4.1 and 4.2 show the details of the protein G C2 domain : Fc and the protein G domain III : Fab interactions respectively.

Protein G, like protein A, can selectively bind to IgGs, but with higher affinity (k_d = ~ 10^{-8} M).[105] So it is also often used to purify IgGs.[221,222] Likewise, high cost limits its application in IgGs purification. Therefore, novel cheap ligands with good selectivity and high affinity are extremely expected.

4.2.1.1 Binding Between Protein G and Fc Regions of IgG

Protein G can bind the Fc fragment very tightly as protein A does. Crystal structure of the protein G C2 domain complexed with the Fc fragment (PDB: 1FCC)[220] of IgG (Figure 4.1a) is known. The binding interface is located between protein G and the hinge region which connects the C_H2 and C_H3 domains of IgG. Three residues, Ile253, Ser254 and Gln311 from the C_H2 domain, and nine residues from the C_H3 domain were involved in the interactions. Of the twelve residues, nine residues are situated within the loop regions of the Fc, three residues: Glu380, Glu382 and Gln438 are exposed on one of several β-sheets making up the C_H3 domain. The residues from the α-helix and the *N*-terminal portion of the third β-sheet of protein G are identified as important residues for binding to the Fc by two independent experiments.[223,224] Sauer-Eriksson's group[220] identified eight residues from protein G as hot-spots: Glu27, Lys28, Lys31, Gln32, Asn35, Asp40, Glu42 and Trp43 (Figure 4.1b).

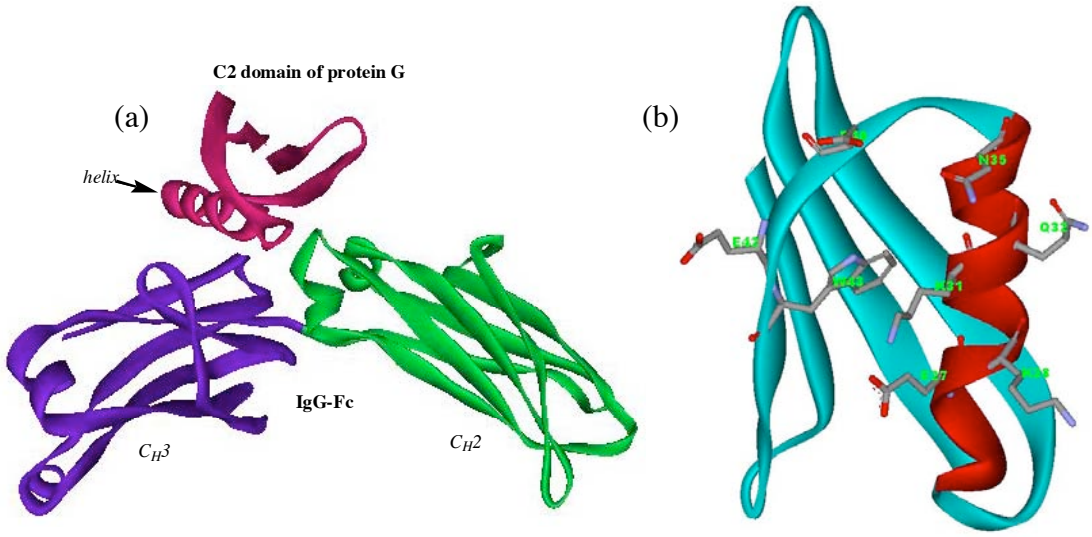


Figure 4.1. (a) The C2 domain of protein G /Fc fragment of IgG complex; (b) hot-spots of C2 domain involved in Fc binding.

4.2.1.2 Binding Between Protein G and Fab Regions of IgG

The interactions between protein G and IgG/Fab feature a sheet region and a helical region on the protein G fragment in close contact with the C_H1 domain of the Fab. Crystal structure of the protein G domain III complexed with the Fab (PDB: 1lGC)[217] of IgG is known (Figure 4.2a). The first of these interactions is an antiparallel alignment of the second β-strand of domain III with the seventh β-strand of the C_H1 domain of Fab (Figure 4.2b). The residues involved in this site are Lys15A to Thr22A from protein G, and Ser209H to Lys216H from the Fab. This β/β interaction site incorporates a few charged residues, Lys18A, Glu20A, Lys212H and Asp214H, which lie along the outer surface of the β-strand and are exposed to the environment. The other, called minor contact, contains residues Tyr38A to Gly43A at the *C*-terminal end of the α-helix from the domain III, and Pro125H to Tyr129H on the first β-strand of C_H1 (Figure 4.2c). These two interaction sites are stacked together so that a huge binding surface area is formed. Hydrogen bonds involved in the interactions are tabulated in Table 4.1.

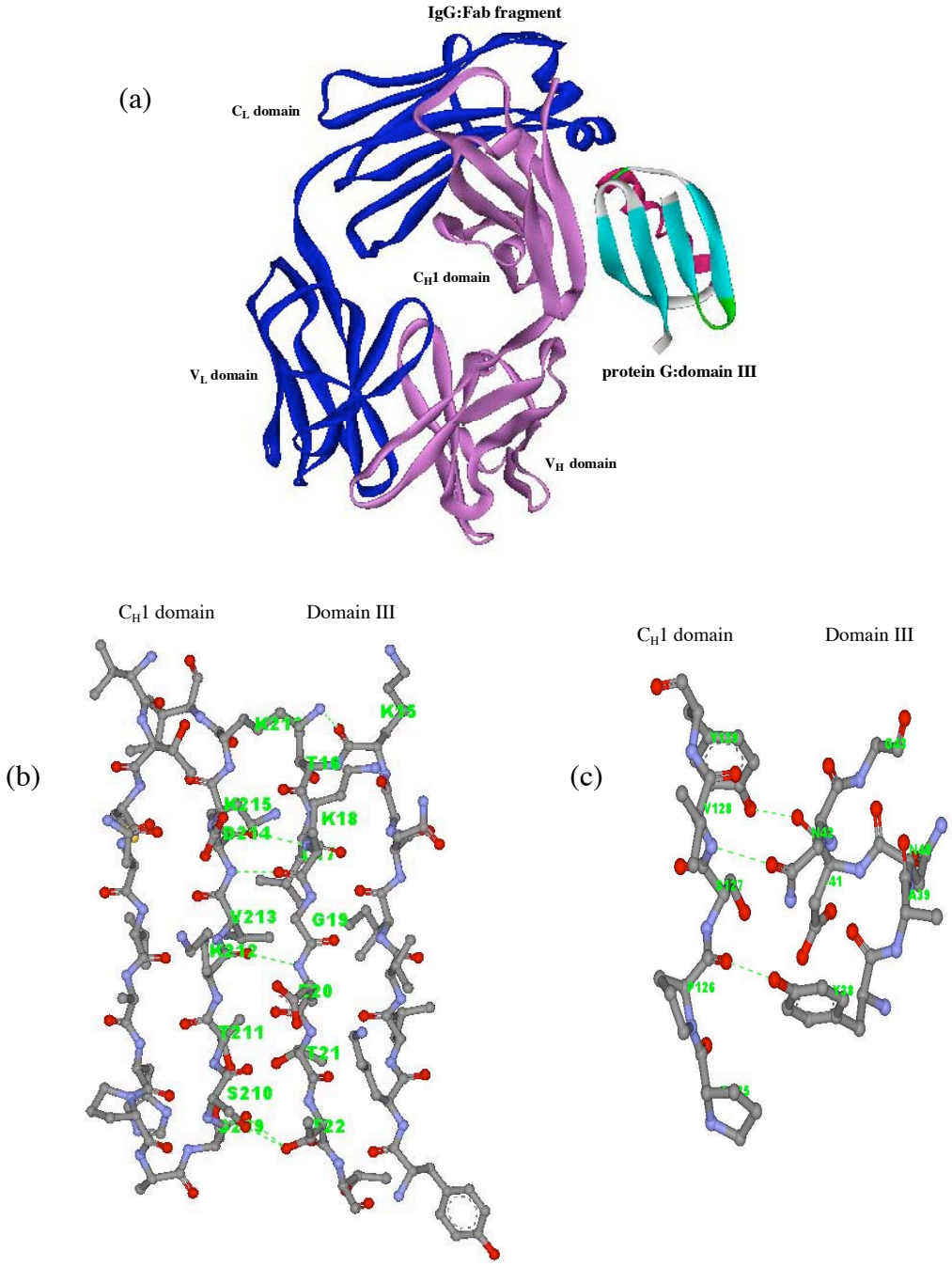


Figure 4.2. ViewerPro plot of interaction details between domain III of protein G and the C_H1 domain of IgG/Fab. Hydrogen bonds between these two domains are marked using dashed lines. (a) The overall interaction; (b) the contact between the Fab and protein G β-sheets; (c) the interaction between a β-strand of the Fab (on the left) and the α-helix of domain III (on the right).

Table 4.1. Hydrogen Bonds Between Domain III of Protein G and the C_H1 Domain of IgG/Fab

donor		acceptor		distance
residue	atom	residue	atom	(Å)
Lys216H	N$^{\zeta}$	Lys15A	O	3.03
Lys18A	N	Asp214H	O	3.18
Asp214H	N	Lys18A	O	2.73
Glu20A	N	Lys212H	O	2.86
Thr22A	OH$^{\gamma}$	Ser209H	O	2.76
Thr22A	OH$^{\gamma}$	Ser210H	O$^{\gamma}$	2.98
Tyr129H	OH	Asp41A	O	2.76
Val128H	N	Asn42A	O$^{\delta}$	3.02
Tyr38A	OH	Pro126H	O	2.96

4.2.2 Comparison of Protein A with Protein G Binding to IgG

Generally, protein G binds all of four human IgG subclasses as well as monoclonal antibodies of mouse and rat. However, protein A binds to neither human IgG3 nor rat IgG.[221,222,225,226]. Protein G binds human IgG with a higher affinity dissociation constant (~10^{-8}M) than protein A (~10^{-7}M) does.[105] Interestingly, both proteins A and G interact with the same part of IgG-Fc,[225,227,228] even though they exhibit no sequence homology within their Fc-binding domains.[108,229,230] The α-helix of either protein A or protein G is involved in the interaction of IgG-Fc. Specifically, The binding between protein A and the Fc is stabilized by hydrophobic interactions and fewer polar contacts, while the binding between protein G and the Fc is stabilized by hydrogen bonds and salt links containing charged and polar residues, and occasional main-chain atoms. Protein A has only 6 charged or polar interactions with Fc, while protein G has 12. However, there are five hydrophobic contacts between protein A and the Fc, while there is no hydrophobic contact between protein G and the Fc. So it is obvious that protein A and G apply different interaction modes for their IgG/Fc binding. Despite these differences, the interesting thing is that several of the Fc residues which interact

with protein A also interact with protein G since the helix 2 of protein A is located in nearly the same region as the third β-strand of protein G.

The complex of protein G with the Fab (Figure 4.2a) displays an entirely different way used in the interaction at the protein-protein interface. The binding between protein G and the Fab is predominantly mediated by the hydrogen-bond formation between main-chain/main-chain, via pairing of two β-strands. This kind of binding is totally different from that with the Fc, where the interactions from side-chain/side-chain predominate, via forming hydrogen bonds and salt links.

4.2.3 Structure Based Design of Protein G Mimics

With regard to protein G mimics, only one paper, " synthesis of a dityrosine-linked protein G peptide dimer" by Vranken's group,[231] was found. Based on the structure of the binding domain (beta/beta strands) between protein G and IgG/Fab, they designed and synthesized the dityrosine-linked peptide dimer using solid phase chemistry (Figure 4.3). The original sequence of protein G B1 domain contributing to the beta/beta interactions with the Fab fragment of IgG is presented on the right, and dashed lines indicate hydrogen bonds. The dityrosine peptide dimer (on the left) was obtained from the replacement of Lys with Tyr and dimerization of the monomer.
In this paper they did not provide evidence for the beta sheet conformation which dityrosine peptide dimer was expected to adopt in solution, and did not present any bioassay data.

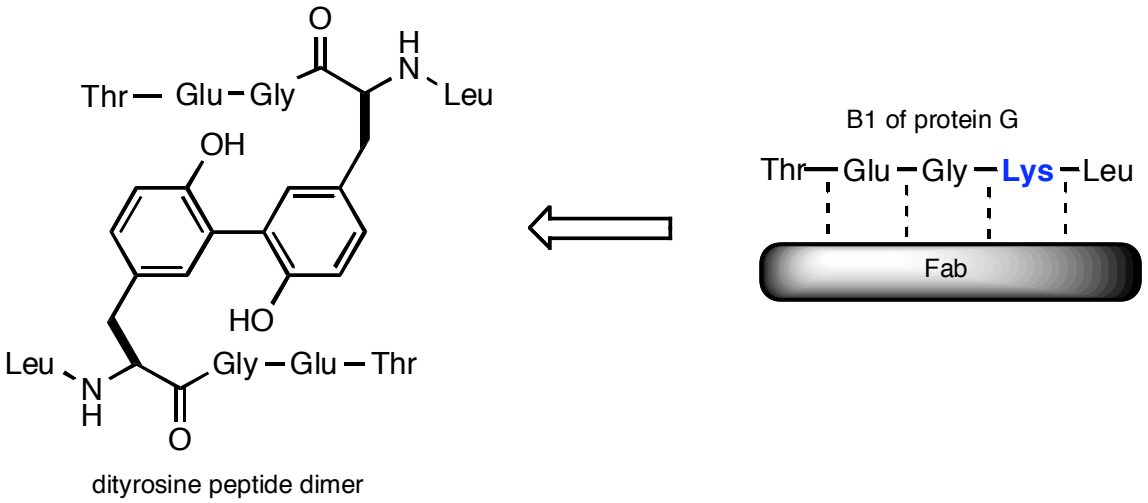


Figure 4.3. Structure of protein G β-sheet/IgG β-sheet interaction mimic.

4.2.4 Functionalized Modification of Poly(propylene imine) Dendrimers

Dendrimers are chemically well-defined, highly branched molecules with a perfect monodisperse and tree-like structures. Well architecture, extensive branching and high surface functionality have distinguished dendrimers from classic polymers. Syntheses, properties and applications of dendrimers are discussed in several reviews and books.[232-239]

Poly(propylene imine) dendrimers are one of the two classes of popularly studied and commercially available dendrimers.[238] Dendric molecules like compound **DAB-Am-4** (Figure 4.4) are attractive frameworks for syntheses of tetravalent and higher oligomers of biologically active ligands. This is important because many protein-protein interactions feature multivalent ligands interacting with receptors to initiate or perturb their association.[2,18,240-243] Within this area of interest there are many situations in which it would be advantageous to attach ligands to all the sites except one on a dendric system; for example, if three out of the four primary amine groups in **DAB-Am-4** were coupled to biomolecules then the remaining unfunctionalized one could support a fluorescent or similar label, or be used to attach it to an affinity support. Further, in

combinatorial chemistry, size exclusion methods may facilitate more facile isolation of dendric molecules than similar monofunctionalized ones.[244,245]

It is extremely difficult to devise reactions to selectively use $n - 1$ functional groups out of a total of n at the periphery of a dendric entity. In fact, to the best of our knowledge, this has not been reported before. The closest literature precedents would be methods in which segments of dendric molecules are functionalized in different ways,[246-248] and these tend to involve *de novo* syntheses of the dendric core.[249] This research attempts to develop an alternative approach, also depicted in Figure 4.4, which takes advantage of the relatively disperse arrangement of functional groups on supports for solid-phase syntheses.[250,251]

The hypothesis that drove this research was that one of several reactive groups could be reacted, selectively, with the resin leaving the others free to be transformed into ligands for biological interactions. This is not the same as other research in which dendrimers have been used to increase the loading of resins because, unlike those studies,[252,253] the goal is to obtain molecules that *contain the dendric core fragment* in the products after cleavage from the resin.

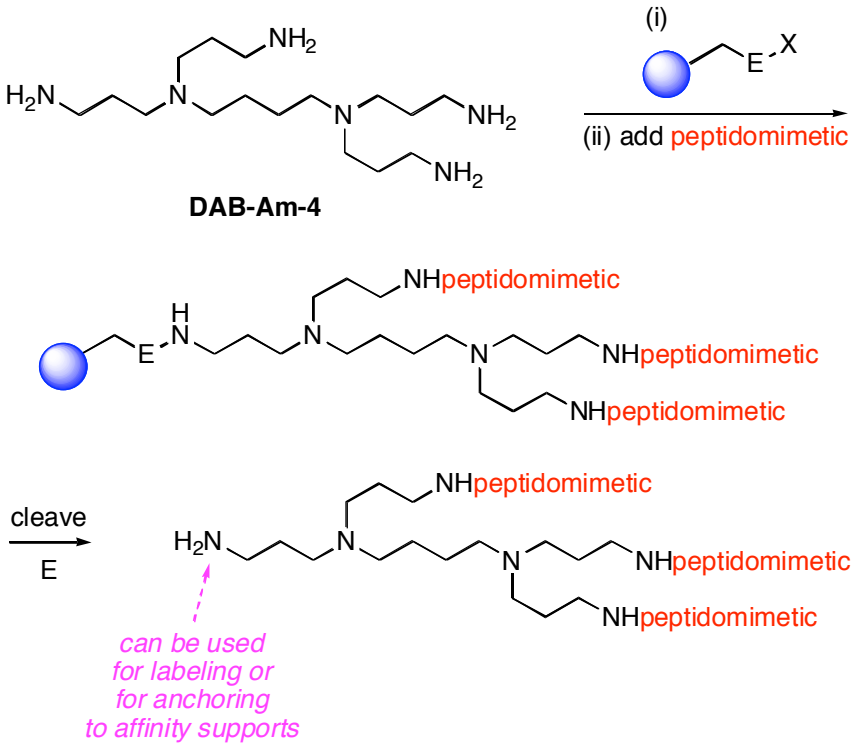


Figure 4.4. Conceptual approach to trifunctionalization of three of the four sites in the tetravalent core of **DAB-Am-4**.

The concept outlined in Figure 4.4 was tested in the context of finding ligands that interact with immunoglobulins G (IgG). The focus of the current research was to find peptidomimetics based on the hot-spots involved in the two interaction sites between protein G/domain III and IgG/Fab (Figure 4.2). We hypothesized that these two interactions form a hydrophilic perimeter around a hydrophobic core formed among Pro125H, Thr211H and Val213H. This is a typical type of hot-spot in protein-protein interactions that has been compared with an "O-ring".[26] We chose to target this region of the IgG Fab fragment for our peptidomimetic design. It was anticipated that one unit on a dendrimer might rest comfortably in the hydophobic pocket while the others could form additional contacts with the protein surface. The site on the dendric molecule that did not contain a peptiomimetic "warhead" would be used to anchor the whole dendric molecule to a solid support.

4.2.5 Affinity: A Software For Docking Ligands to the Binding Sites

Affinity is a suite of programs which is designed for automatically docking a ligand molecule to a receptor molecule. For a given system of a ligand and a receptor, Affinity can automatically find the properly binding structure for this system based on the energy of the system on Insight II platform.

There are a few features in Affinity. First, Monte Carlo[254] and Simulated Annealing strategy[255], both of which are powerful algorithms, were employed to dock a ligand to a receptor. Second, a full molecular mechanics forcefield was used in searching for and evaluating docked structures. That is, Affinity takes into account electrostatic factors, hydrogen bonding, hydrophobic interaction and steric effect. Third, a receptor may be divided into two sections: one is the binding site, in which atoms are active and move freely; the other is called "bulk", in which atoms are not in the binding site and are held rigidly during the docking process. This feature allows a receptor to adjust either the binding of different ligands or different binding modes of the same ligand. Fourth, the ligand itself is flexible, which allows different conformations of the same ligand to be docked to a receptor. Fifth, a three dimension grid[256] surrounding the movable atoms containing the binding site atoms of a receptor and all atoms of a ligand may be constructed and presents as a "bulk" in order to reach the above aims. The following steps are required for docking ligands to the binding site: **a**. download a receptor from protein data bank and fix incorrect part of the structure; **b**. build a ligand using Builder module in Insight II; **c**. define a subset from a receptor for the binding site; **d**. create an assembly of a ligand and a receptor; **e**. assign potentials for the assembly; **f**. generate grids; **g**. perform docking calculation; **h**. get docked structures and analyze results.

4.3 Design of Protein G Mimics for IgGs

The Affinity module of Insight II was used to explore the docking of several molecules that appeared to have desirable features for binding to the IgG Fab region.

Affinity works in the following way. The user defines a region of the target protein surface and a synthetic ligand, then the docking process is simulated using the Affinity module. It minimizes the conformation of the ligand *and* the user-defined region of the protein surface hence accommodating the possibility of an induced fit. The particular region of IgG/Fab considered was the one that encompasses the hot-spots for binding of IgG to protein G domain III, *ie* a ~210 $Å^2$ area on the IgG/Fab surface between the two interaction sites shown in Figure 4.2b and c. This region of the protein surface is essentially flat, but it contains a hydrophobic channel and some Lewis basic site (formed by Ser127H, Ser209H and Lys215H amino acids of the Fab fragment). We hypothesized that the hydrophobic channel might accommodate a residue such as Leu17A, Tyr38A and/or Thr21A on protein G, while the H-bond acceptors might match with the Lys15A, Lys18A, Glu20A, Thr22, or Asn42A of protein G. Synthetic structures that might bind were selected for compatibility with these surface characteristics, and their synthetic accessibility, then their binding was tested using Affinity. Ultimately this led to compounds **A – C** that gave favorable binding modes in the Affinity-simulated docking to the selected patch of the IgG surface. Figure 4.5 shows their simulated binding modes.

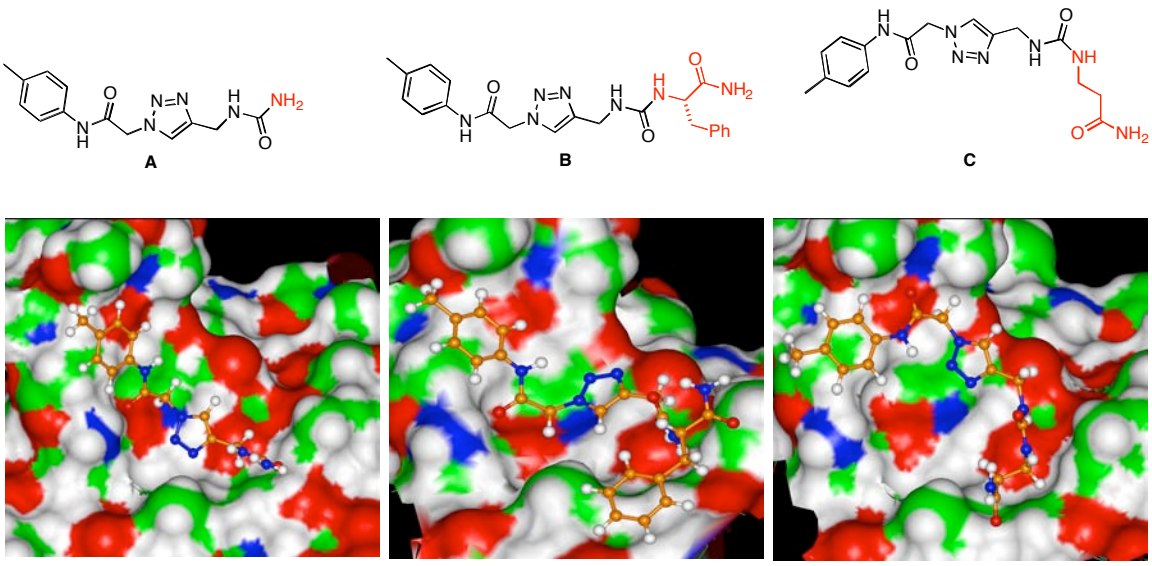


Figure 4.5. Docking of the virtual leads **A** – **C** with the region of IgG that docks with the protein G domain III.

Structures **A** – **C** were treated as "virtual leads", from which a library of 18 trivalent compounds was designed, focused on these structures (Figure 4.6). The concept was that a supported form of the **DAB-Am-4** polyamine would be coupled with the three alkyne-electrophiles, or electrophile precursors, **8a-c**, and the products would be combined with the seven azides **9t** – **z** to give products like **10**.

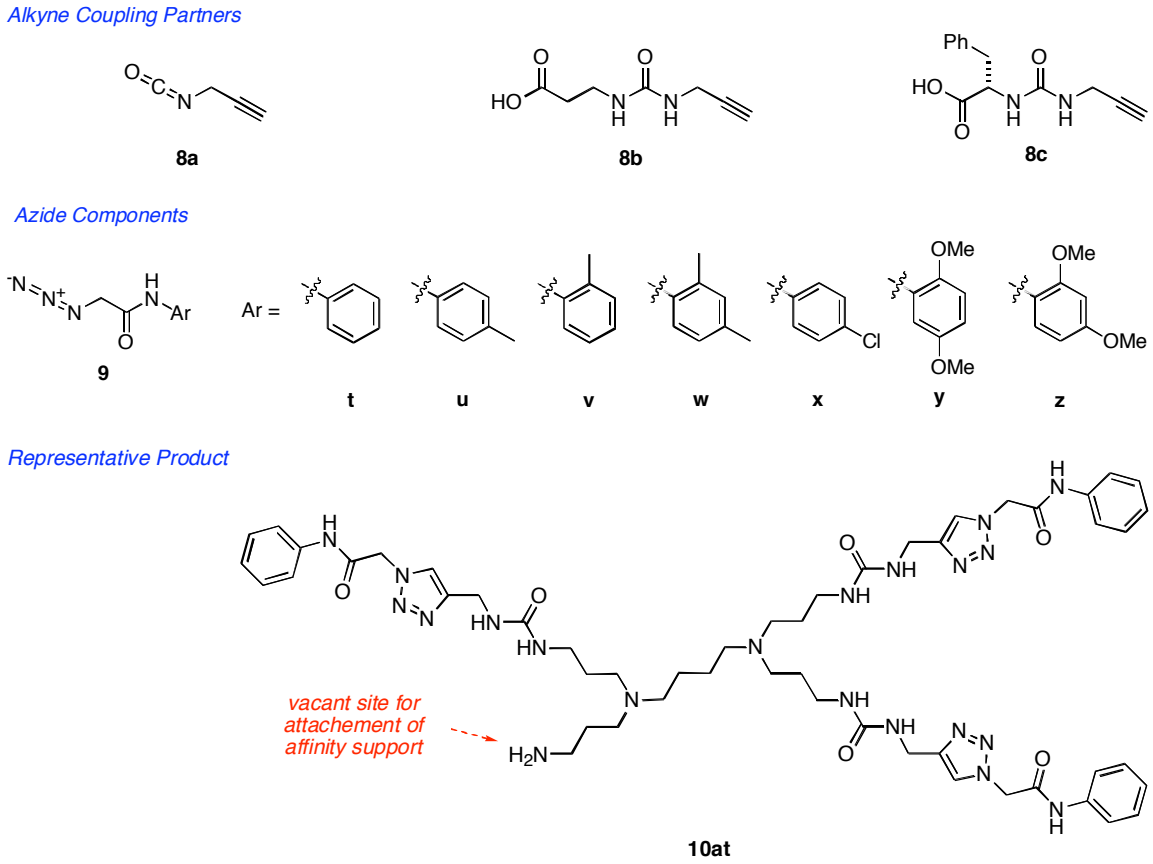


Figure 4.6. Library design: it was anticipated that the alkynes **8a-c** would be coupled a supported dendric amine, then with the azide components **9** to give the products like **10**.

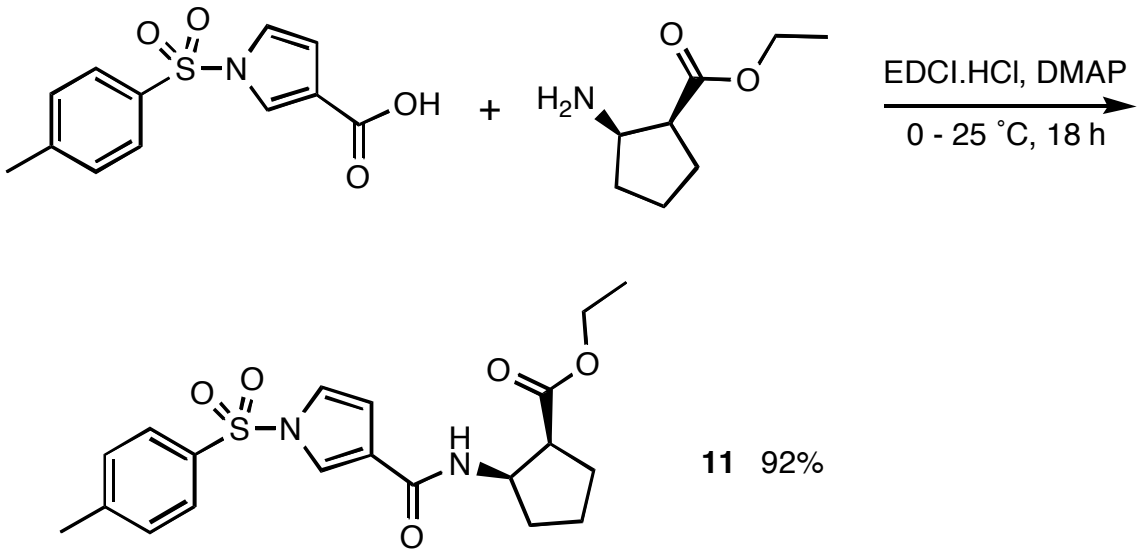
4.4 Syntheses of Monovalent Protein G Mimics, Functional Compounds and Building Blocks for Multivalent Protein G Mimics

4.4.1 Synthesis of Compound 11

Scheme 4.1 describes synthesis of compound **11** which was designed and proven to be active to the binding of IgGs by Amersham. This is a one-step reaction. Compound **11** was obtained by stiring the mixture of *N*-tosyl-3-pyrrolecarboxylic acid and ethyl-cis-2-amino-1-cyclopentane carboxylate for over ten hours. The Carboxy group of *N*-tosyl-3-pyrrolecarboxylic acid is conjugated to the pyrrole ring, so several coupling agents were tried for this coupling reaction: (a) DIC and HOBt, (b) oxalyl

chloride, (c) DIC or EDCI and DMAP, (d) TFFH. TLC indicated that (b) and (c) work well, and compound **11** was obtained in 92% yield.

Scheme 4.1. Preparation of Compound 11

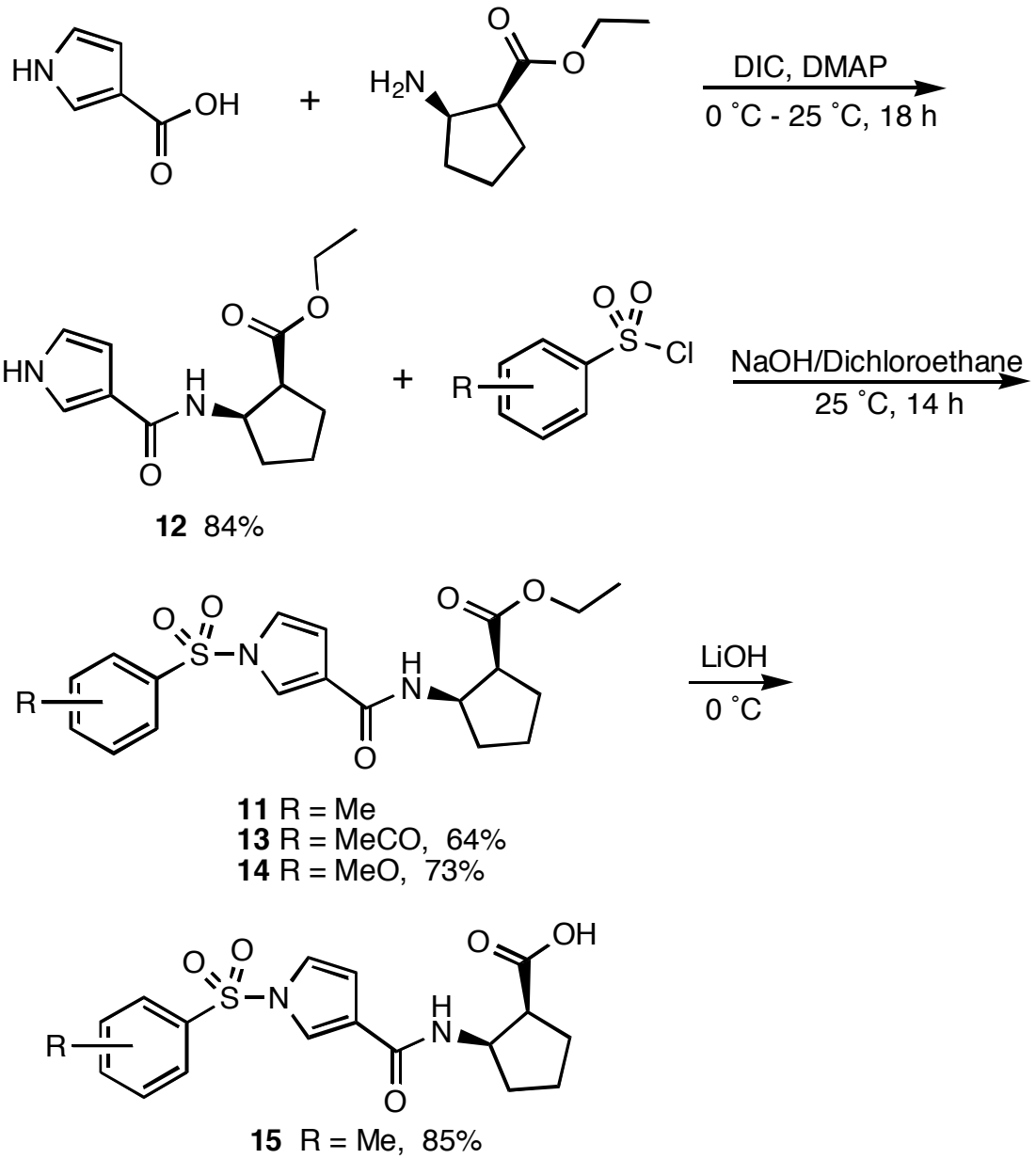


4.4.2 Syntheses of Functional Compounds for Multivalent Protein G Mimics

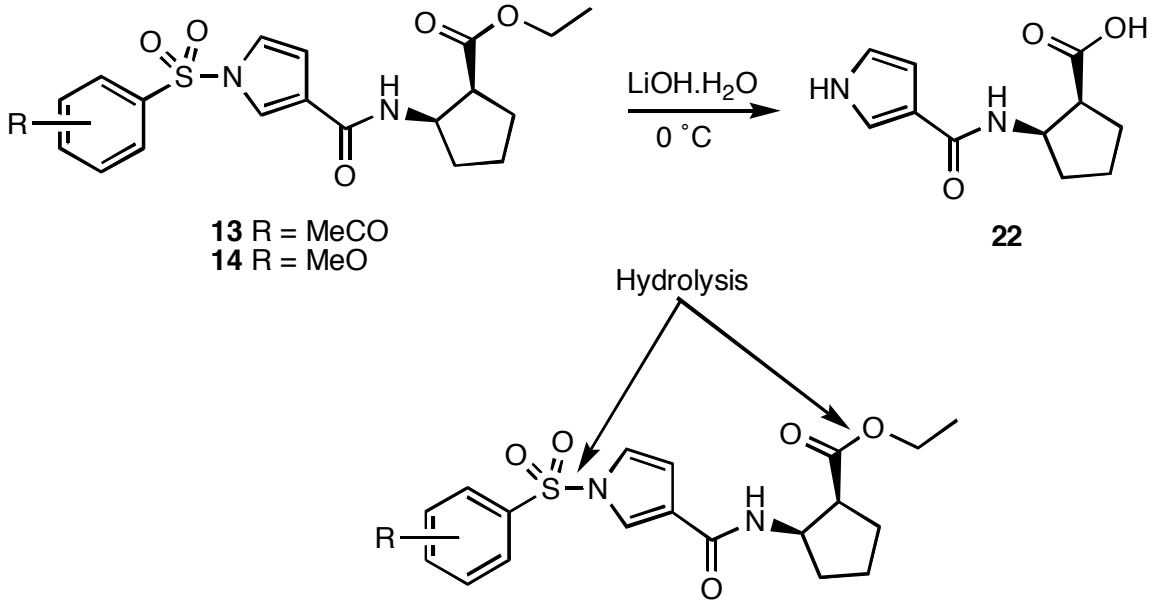
Two approaches were used to prepare functional compounds. Scheme 4.2 depicts approach *I* applied to synthesize functional compound **15** (R = Me). The coupling between pyrrole-3-carboxylic acid and ethyl-cis-2-amino-1-cyclopentane carboxylate under the conditions of DIC or EDCI and DMAP gave intermediate **12** in 84% yield, which reacted with substituted benzene sulfonyl chloride, and afforded

compound **13** in 64% yield and **14** in 73% yield. By carefully controlling the temperature and reaction time, the functional compound **15** was obtained in 85% yield. This method was also used to synthesize other functional compounds (R = 4-MeCO, 4-MeO). Unfortunately, the hydrolysis of compounds **13** and **14** does not work well. The yields are very low, and the major product is compound **22** (Scheme 4.3). Therefore, approach *II* was employed to prepare these functional compounds (Scheme 4.4). Intermediate **12** was synthesized as described in approach I, and hydrolyzed completely to give intermediate **22** in 90% yield, which was followed by the substitution[257] of the *N*-proton of the pyrrole ring with substituted benzene sulfonyl group and gave the final functional compounds **16 - 21** (Table 4.2). In approach *II*, we found that if the substituted group in benzene ring of aromatic sulfonyl chloride is an electron-donating group, the substitution works well and gives functional compound in higher yield. If the substituted group is an electron-withdrawing group, the substitution does not work well. The position of the substituted group in the benzene ring also affects the reaction. The substituted groups in the *ortho* and *p*ara positions of the benzene ring facilitate the reaction, while the substituted groups in the *meta* position of the benzene ring is against this reaction.

Scheme 4.2. Synthesis of Functional Compounds (*approach I*)



Scheme 4.3. Hydrolyses of Compounds 13 and 14



Scheme 4.4. Synthesis of Functional Compounds 16-21 (*approach II*)

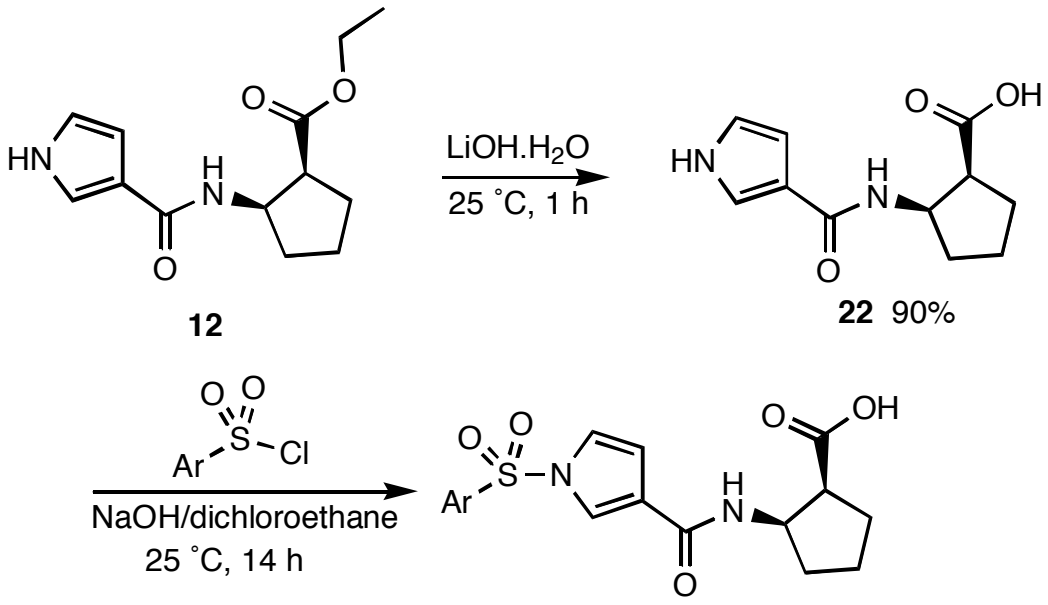
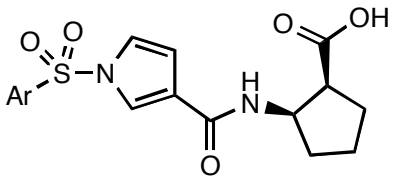
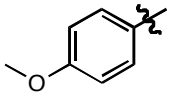
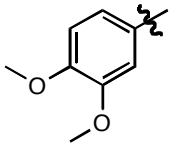
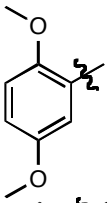
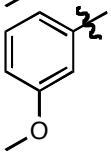
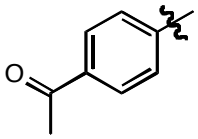
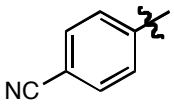


Table 4.2. Some Functional Compounds



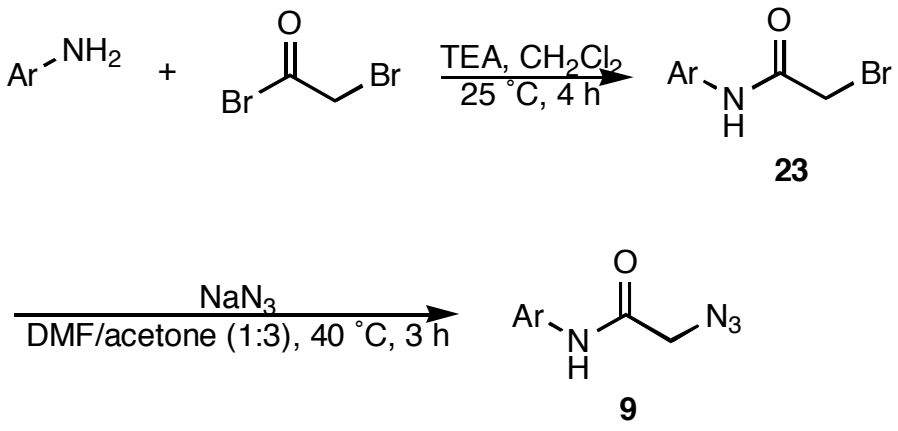
compound	Ar	yield (%)
16		76
17		77
18		69
19		51
20		a little by TLC
21		a little by TLC

4.4.3 Syntheses of Building Blocks for Multivalent Protein G Mimic Library

One building block is 2-azide-*N*-(aryl)-acetamide **9**, which are prepared via two-step reactions (Scheme 4.5). First, substituted aniline reacts with bromoacetyl bromide under the condition of triethyl amine in an ice bath.[258] This reaction goes fast and affords 2-bromo-*N*-(aryl)-acetamide **23** in about 90% yield. To decrease side-products and let the reaction go steadily, the slow addition of bromoacetyl bromide into the solution of substituted aniline and lower temperature are necessary to this reaction. All 2-bromo-*N*-(aryl)-acetamides **23** were purified via recystallization before the next reaction. The substitution in compounds **23** with sodium azide goes smoothly at 40 ˚C in a few hours, and gives 2-azide-*N*-(aryl)-acetamide **9** in high yields[259] (Table 4.3). The use of co-solvents of acetone and DMF is for easy work-up after the end of this reaction.

The other building block is alkyne derivatives **8a-c (**Scheme 4.6**)**. Compounds **8b** and **8c** were synthesized using the same approach. Amino acid methyl ester was treated with triphosgene and saturated aqueous sodium hydrogencarbonate for 20 min.[260,261] The reaction gave amino acid ester isocyanate **24**. To afford high yield, lower temperature and vigorous stirring for this reaction are important. Other methods were also tried, such as the use of DIEA instead of sodium hydrogencarbonate. However, the yield of amino acid ester isocyanate is very low. To amino acid ester isocyanate **24** in CH_2Cl_2 was added propargylamine. After 1 h the reaction was worked up. The crude material was then columned quickly and gave compounds **25** in 94% yield and **26** in 95% yield, which was hydrolyzed with lithium hydroxide monohydrate, and the final product with a carboxylic acid group was obtained. Compound **8a** was synthesized using the same method as for amino acid ester isocyanate **24**.

Scheme 4.5. Syntheses of 2-Azide-*N*-(Aryl)-Acetamide 9



Scheme 4.6. Syntheses of Building Blocks 8a-c

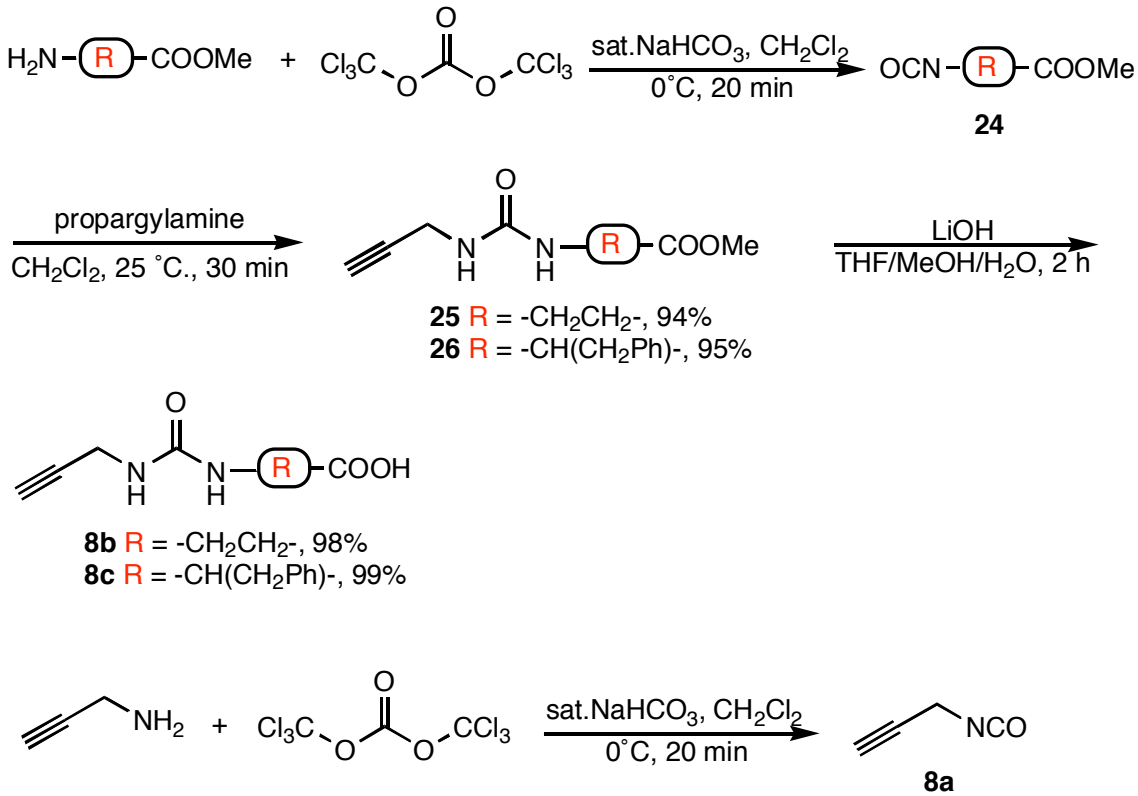
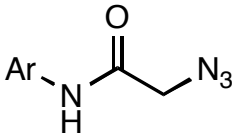
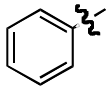
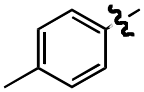
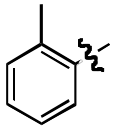
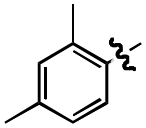
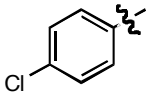
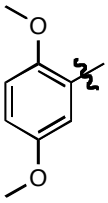
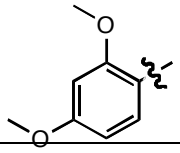


Table 4.3. 2-Azide-*N*-(Aryl)-Acetamide



9

compound	Ar	yield (%)
9t		85
9u		90
9v		90
9w		93
9x		88
9y		89
9z		87

4.5 Development of Solid Phase Syntheses of Multivalent Protein G Mimics Using Poly(propylene imine) Dendrimers as Scaffolds

4.5.1 Choice of Resins and Linkers, and Loading Determination of Resins

For any solid phase synthesis, the selection of a resin is important. Many kinds of resins are commercially available. Commonly, a suitable resin may be selected based on reaction group, substrate, original loading, cleavage conditions and so on. Sometimes an experiment must be carried out in order to confirm if a resin is suitable for the specific synthesis. Bradley and Gawley reported that PS-PEG (polystyrene-poly(ethylene glycol)), TentaGel and ArgoGel could be used in dendrimer synthesis.[250,251] In this study several resins were tested: TentaGel S PHB, NovaSyn TG Hydroxy, Wang and PEGA resins. The loadings (Table 4.4) of these resins except for PEGA were determined by standard quantitative detection of the liberated FMOC byproduct via UV spectroscopy.[262]

Table 4.4. Original and Estimated Loading Values for Several Resins

resin	original loading (mmol/g)	estimated loading (mmol/g)	DAB-Am-n
TentaGel S PHB	0.24	0.227 ± 0.011	n=4
TentaGel S PHB	0.24	0.28 ± 0.03	n=8
NovaSyn TG Hydroxy	0.3	0.22 ± 0.012	n=4
Wang	1.3	0.54 ± 0.025	n=4

Even though Wang resin is cheap, the estimated loading is as high as 0.55 mmol/g. It is not suitable to be used for the modification of dendrimers, especially high generation dendrimers. After consideration of a hard cleavage from NovaSyn TG Hydroxy and poor results from PEGA resin, TentaGel S PHB was finally chosen to be used in the syntheses of multivalent protein G mimics.

To attach poly(propylene imine) dendrimer to the solid matrix, a suitable linker was required. Several linkers[263-265] are shown in Figure 4.7. Linkers **a** and **b** were used for a model study. It was found that **b** is better than **a**, so **b** was selected as the linker for multivalent protein G mimic syntheses.

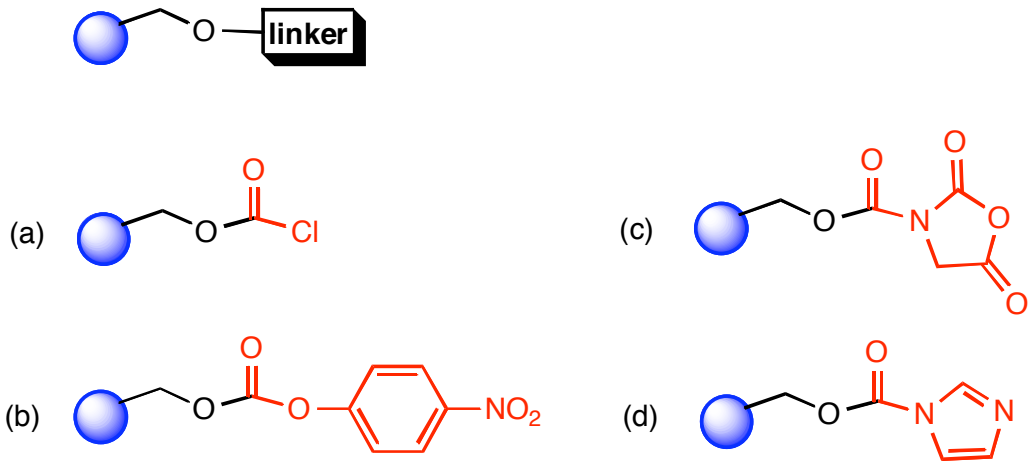


Figure 4.7. Several linkers for the syntheses of multivalent protein G mimics.

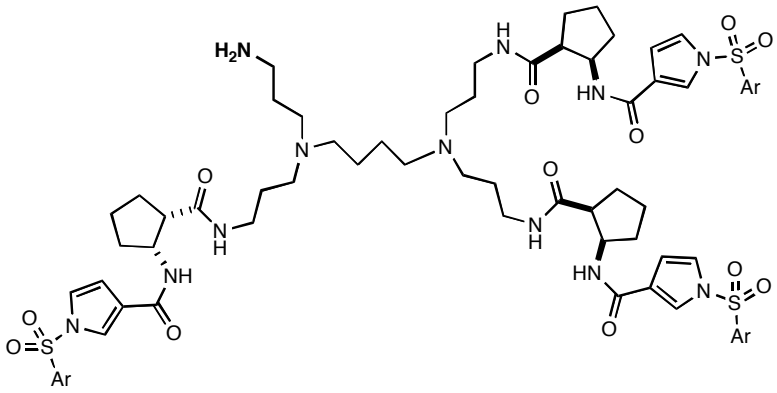
4.5.2 Syntheses of Multivalent Protein G Mimics Based on Amersham-Designed Compounds and My Peptidomimetic 6 Discussed in Chapter III

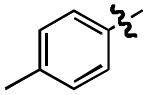
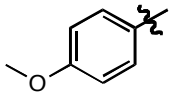
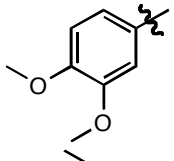
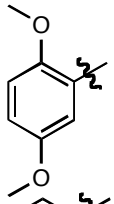
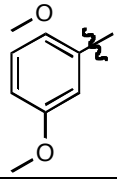
Schemes 4.7 and 4.8 outline the strategies that were used to obtain multivalent protein G mimics based on DAB-Am-4 (generation 1) and DAB-Am-8 (generation 2) dendrimers respectively. The swelled TentaGel S PHB was treated with 4-nitrophenylchloroformate and *N*-methyl morpholine,[264] and followed by the treatment of poly(propylene imine) dendrimers and DMAP. After the mixture was gently shaken for two days, the free amine resin-bound intermediate **28** or **30** was obtained. In this step, the ninhydrin test was carried out, and showed obviously positive results. In terms of different generations of DAB dendrimers, the different ratios of DMF and DCM should be used for this reaction. Amide formation between the intermediate **28** or **30** and functional compounds with the carboxylic acid group was carried out for about one day by using standard carbodiimide coupling conditions, such as HOBt and HBTU, HOBt and DIC, and by transforming the carboxylic acid group of functional compounds into carboxylic chloride. Consequently, HOBt and HBTU are the best coupling agents for this reaction.

Finally, the treatment of resin-bound functionalized DAB dendrimers with 50% TFA afforded functionalized DAB dendrimers **29** or **31** with one free amine arm. Table 4.5 shows purity and yield data for functionalized DAB dendrimers **29**. The starting materials of functional compound are expensive, so only a little excess functional compound was used in the amide formation step. To compare the binding activity of peptidomimetic **6** (see Chapter III) with its dendric molecule, compound **32** was synthesized (Scheme 4.9).

All multivalent protein G mimics gave satisfactory molecular ions in MALDI-MS analyses. However, MS spectra of the crude products showed traces of other side products that were formed by either one or two amine-capped DAB-Am-4 dendrimers with functional compounds. Even though one reason for this is that only a little excess of functional compounds was used in the reaction, the key issue was the concentration of poly(propylene imine) dendrimers in the coupling with p-nitrophenylchloromate activated carbonate, and this will be discussed later. To characterize the products, only 1D NMR spectra can't satisfy this. Compound **29a** was taken as an example. Its TOCSY spectrum was analyzed in Figure 4.8.

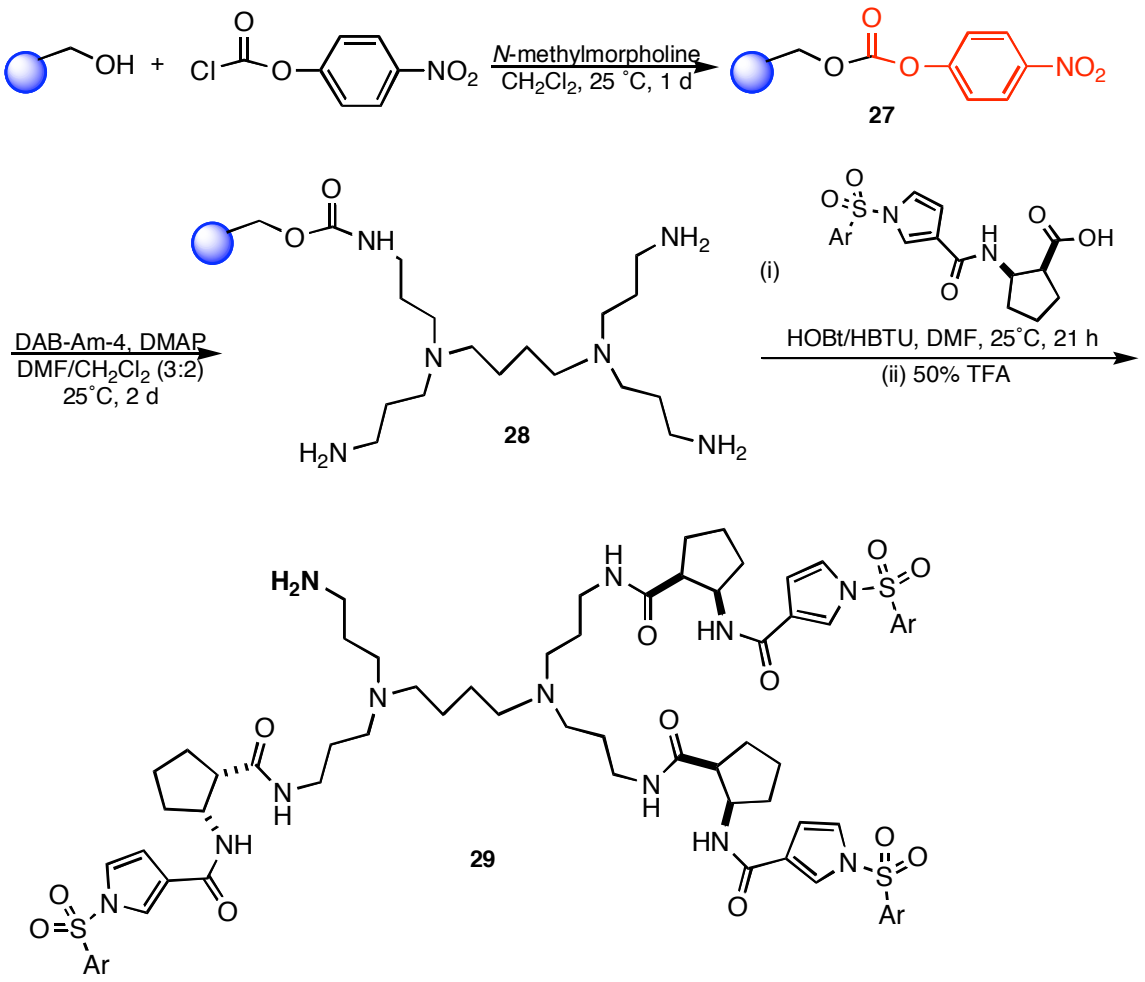
Table 4.5. Summary of Yield and Purity Data for Compounds 29



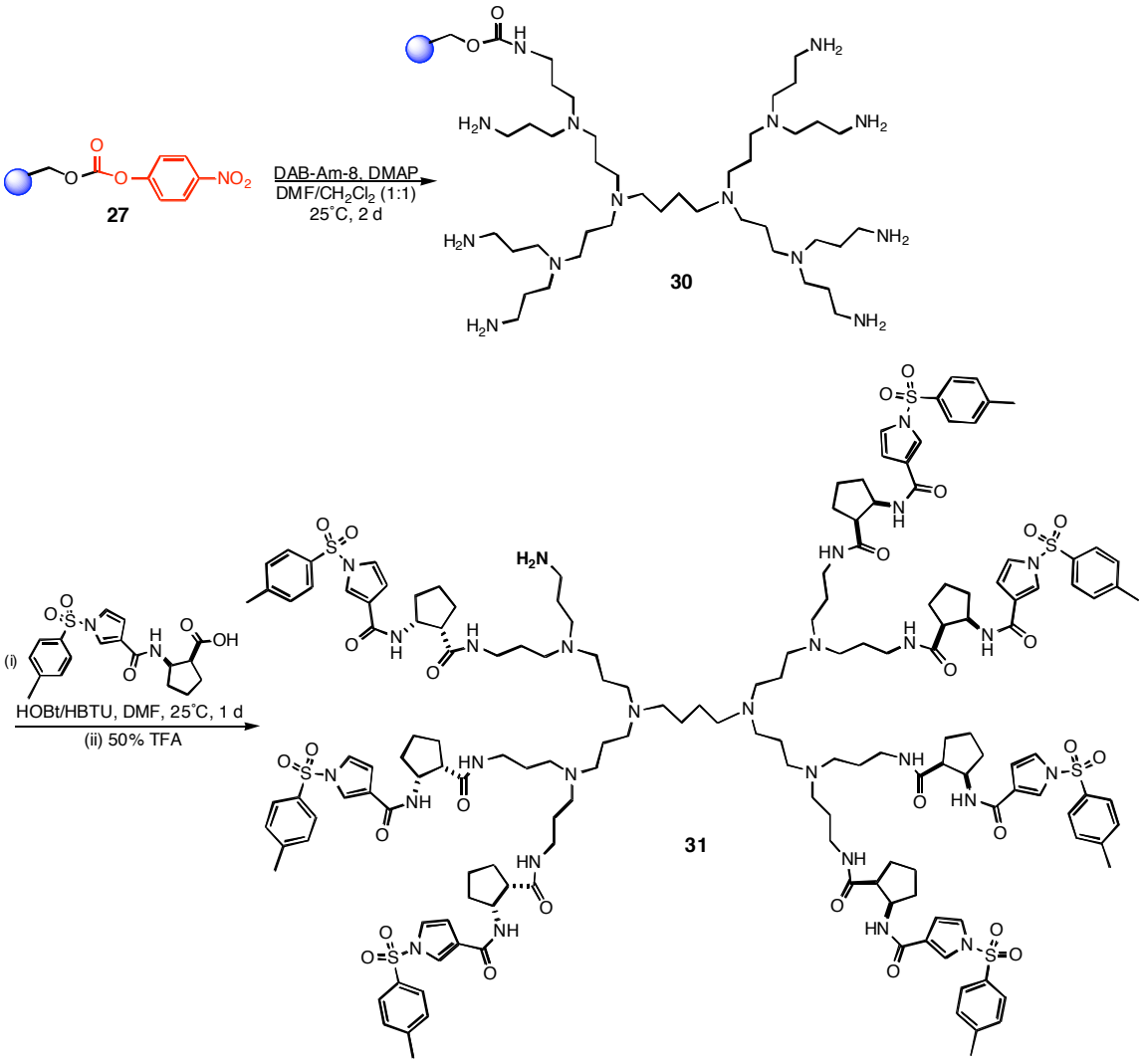
compound	Ar	purity[a] (%)	yield[b] (%)
29a		54	36
29b		76	58
29c		82	49
29d		74	45
29e		62	33

[a]Purity was assessed by HPLC for crude product via monitoring UV absorption at 254 nm. [b]Yield was calculated according to isolated product.

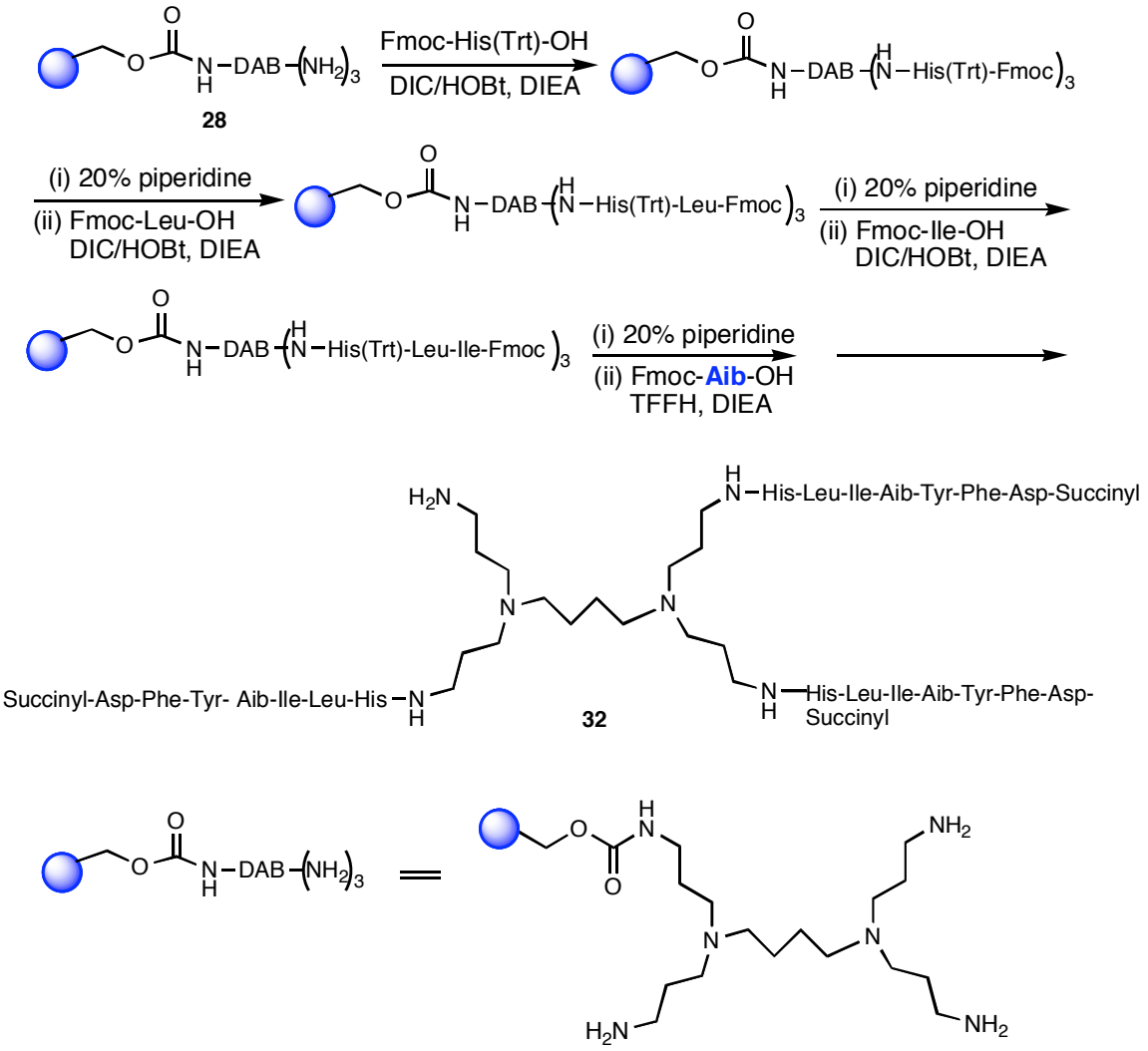
Scheme 4.7. Preparation of Trivalent Compounds 29



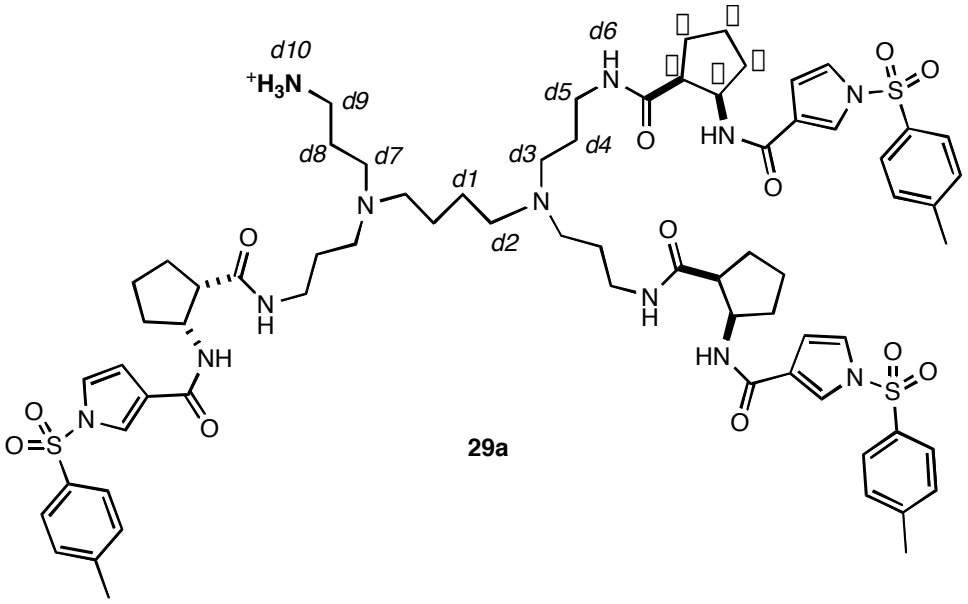
Scheme 4.8. Preparation of Multivalent Compound 31



Scheme 4.9. Preparation of Trivalent Compound 32



(a)



(b)

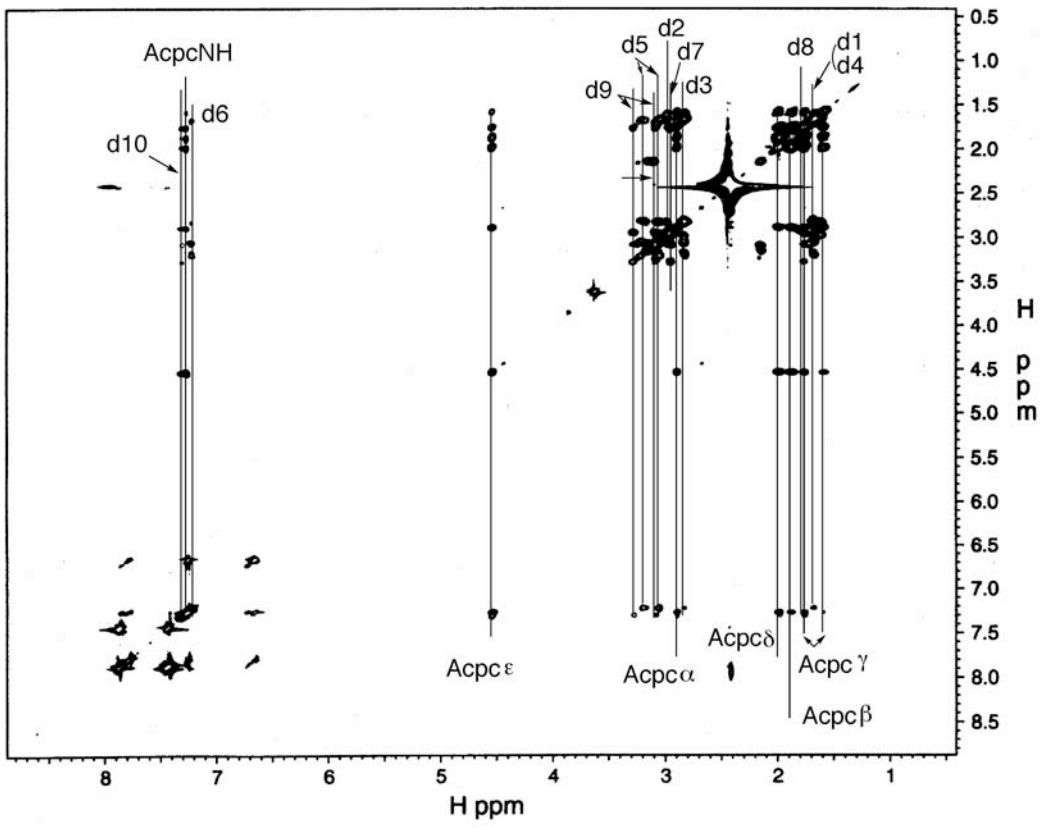


Figure 4.8. Structure (a) and TOCSY spectrum (b) of compound **29a**.

4.5.3 Library of Trivalent Protein G Mimics Based on Our Designed Monovalent Ligands

Even though some reaction conditions of the strategy used in Scheme 4.7 has been optimized, the crude products of **29** still have poor purity. The same strategy as described in Scheme 4.7 was used to synthesize multivalent protein G mimics using our designed monovalent ligands. One and two amine-capped DAB-Am-4 dendrimers with these ligands are also major byproducts although 5 fold reagents were used in the coupling reactions. After careful evaluation of each step in the synthesis strategy, it became evident that the efficiency of the synthesis with respect to formation of the desired product **10bu** was highly dependent on the concentration of poly(propylene imine) dendrimers used in its coupling with p-nitrophenylchloromate activated carbonate. To optimize this reaction (from **27** to **33** in Scheme 4.10), different concentrations of the **DAB-Am-4**: 0.08 M, 0.15 M, 0.30 M, 0.58 M, 1.0 M, 2.0 M and neat were used while all other conditions used were the same. Figure 4.9 shows the ratios of the three products with varying concentrations of the dendric amine. The optimal concentration with respect to formation of **10bu** was 2 M; if higher concentrations were used then the reaction liquid phase becomes quite viscous and the coupling efficiency declines. Indeed, washing steps in the synthesis became difficult when 2 M concentrations of the dendric amine were used; consequently 1.5 M was preferred.

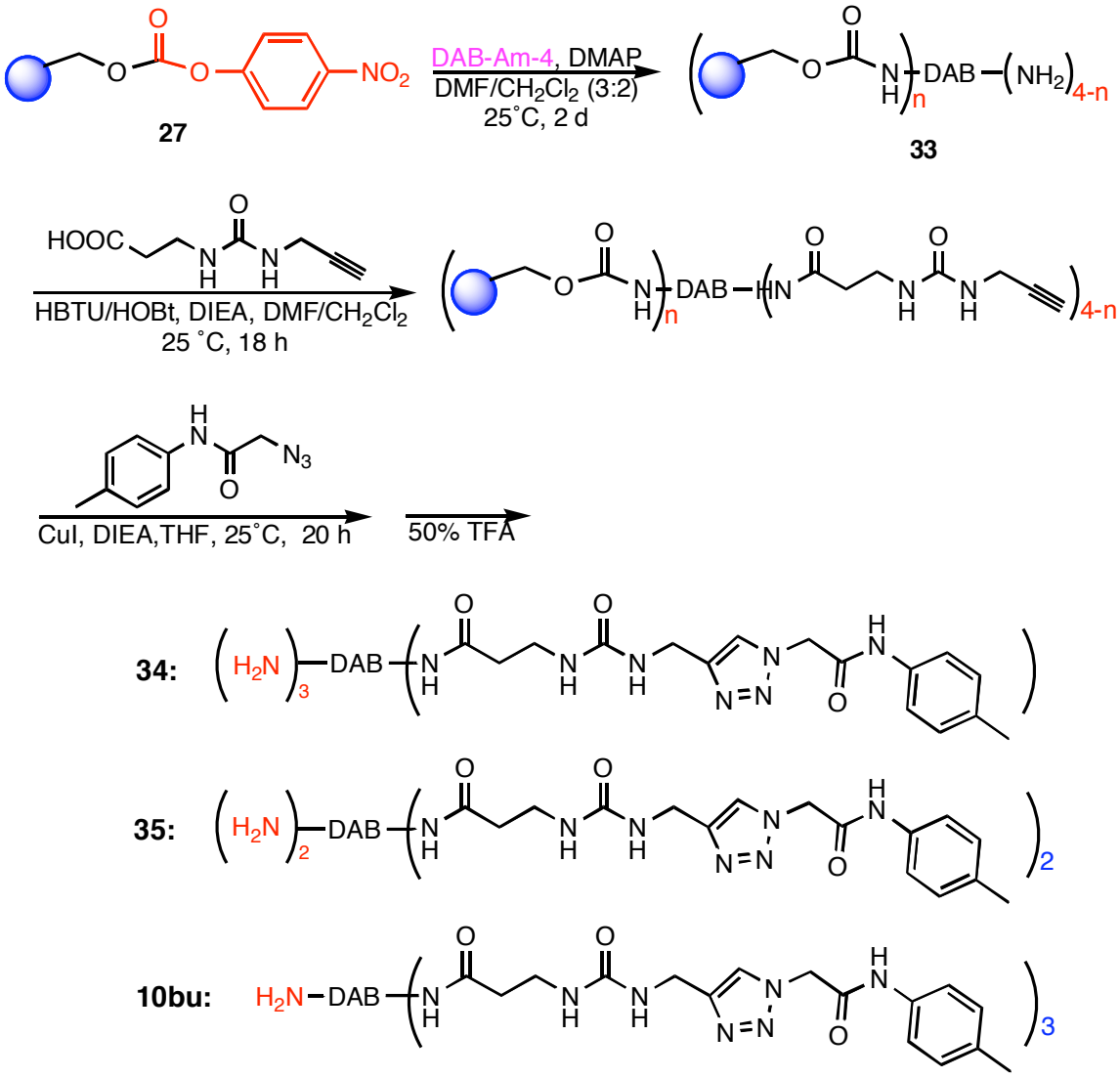
In addition, 1,3-dipolar cycloaddition between alkyne and azide groups on resin was applied in the synthetic strategy (Schemes 4.10 and 4.11). This is a typical click reaction done by a few research groups using either solid phase or solution phase.[266-270] Based on the conditions used in these papers the 1,3-dipolar cycloaddition on resin was also optimized. The tested conditions include different equivalents of copper iodide: 0.1, 1, 2 and 3 equiv, different reaction times: 20, 48 and 72 h, and different temperatures: 25 ˚C and 40 ˚C. Except for the results in which 0.1 equiv. of copper iodide was used, not much difference was found for all other conditions. Therefore, in the library syntheses,

2.5 equiv. of copper iodide, 5 equiv. of azide and 25 ˚C were used for the 1,3-dipolar cycloaddition (Scheme 4.11).

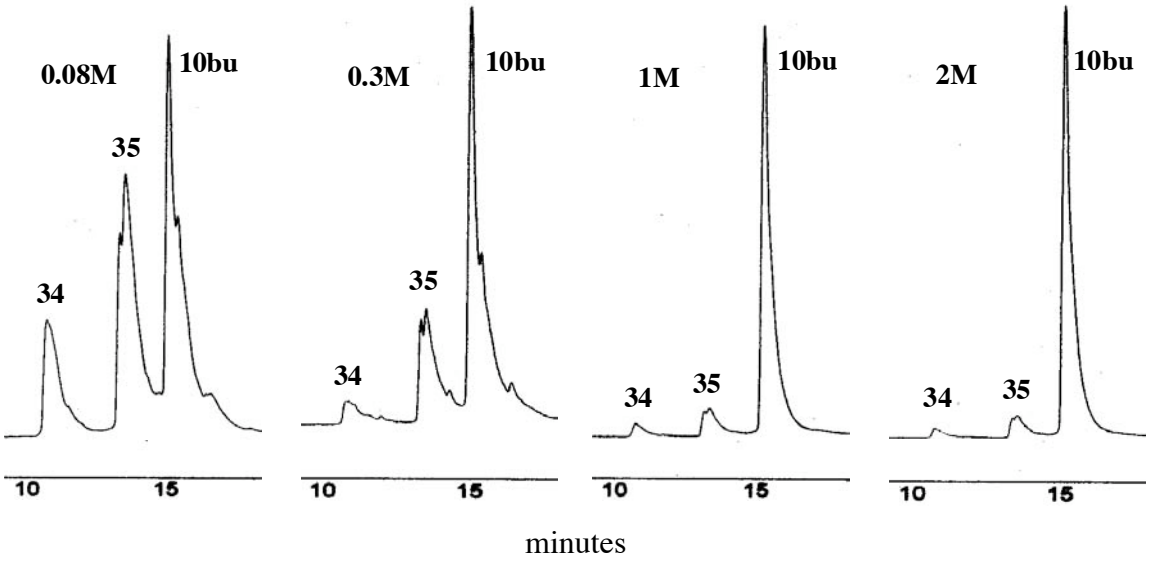
The targeted library of 18 compounds were prepared using the conditions developed above. Table 4.6 outlines details of the purities of the materials cleaved from the resin, and the yields of the desired products based on the loading of the resin. Overall, the data is very satisfactory. The purities of most compounds are greater than 85% (UV) and 96% (Sedex) except that those of compounds **10ay** and **10az** are less than 85%. The reason is that alkyne **8a** was kept over two weeks at room temperature, and not used freshly to make these two compounds. Generally, the purities measured by the Sedex detector are higher than those from a UV detector since different thresholds of detection may be used for these two detectors. In this study, the masses of all 18 compounds are greater than 1000; Sedex, being based on molecular mass, is likely to be more representative of the real purity. Average purity of the crude products was 88.6 % based on UV detection, and 95.8 % based on Sedex. 10 compounds were purified via reverse phase HPLC and characterized via MALDI-MS and ^{1}H NMR. Selected compounds were further characterized via ^{13}C NMR and ^{1}H-TOCSY. TOCSY proved to be particularly informative. Figure 4.10 shows an illustrative spectrum wherein a primary ammonium resonance (D10 in Figure 4.10) correlates with one arm of the dendric backbone whereas the peaks in the other spin systems are largely coincident.

The above optimized conditions were also applied for the syntheses of multivalent protein G mimics using DAB-Am-8 (8 amino groups) and DAB-Am-16 (16 amino groups) as scaffolds. For the use of DAB-Am-8, the crude purity is about 65% (Sedex) while, for DAB-AM-16, HPLC analyses indicated a mixture of products were formed and it proved too difficult to separate the desired products from the other materials (data not shown).

Scheme 4.10. DAB-Am-4 Concentration Dependence Experiments of Products



(a)



(b)

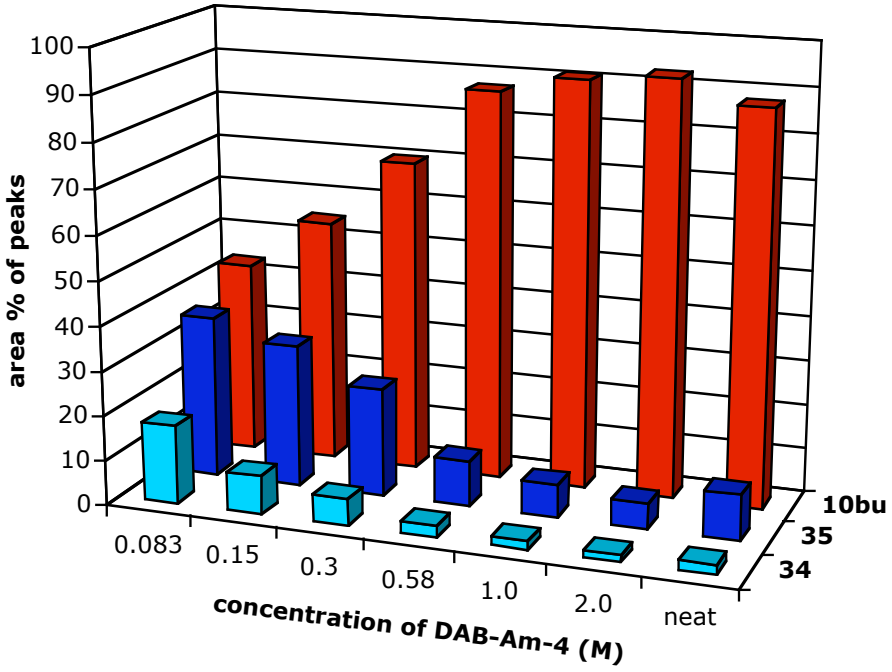
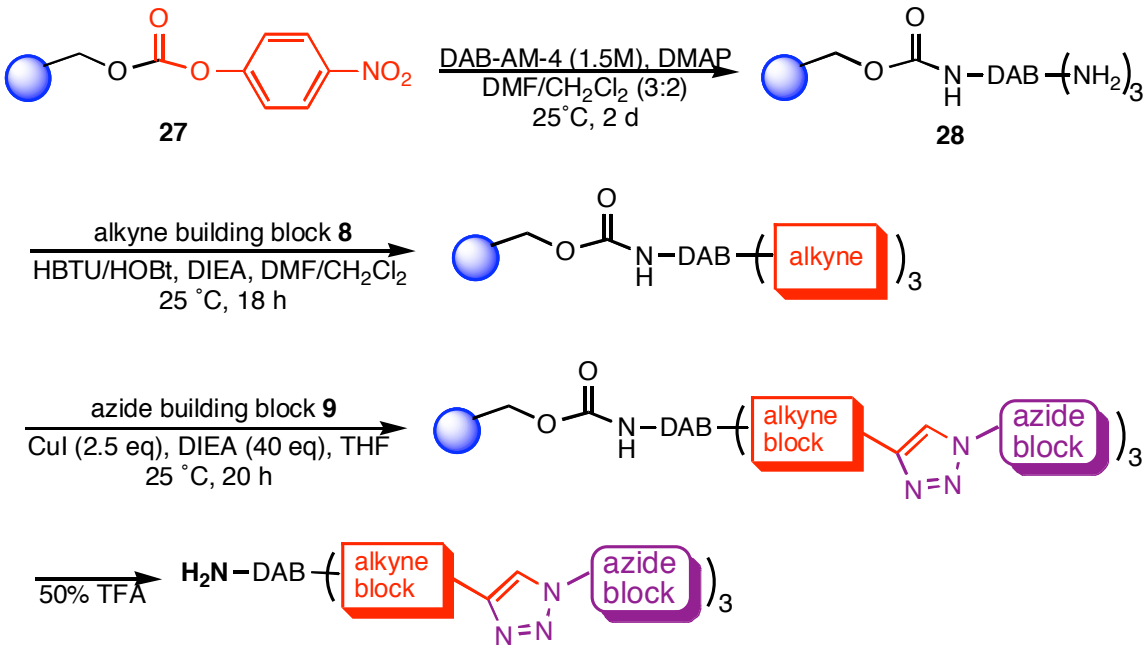


Figure 4.9. (a) Selected HPLC traces indicating the product distribution of the desired product **10bu** relative to the impurities **34** and **35**; and (b) comprehensive data from this set of optimization experiments.

Scheme 4.11. Preparation of Trivalent Protein G Mimics Library



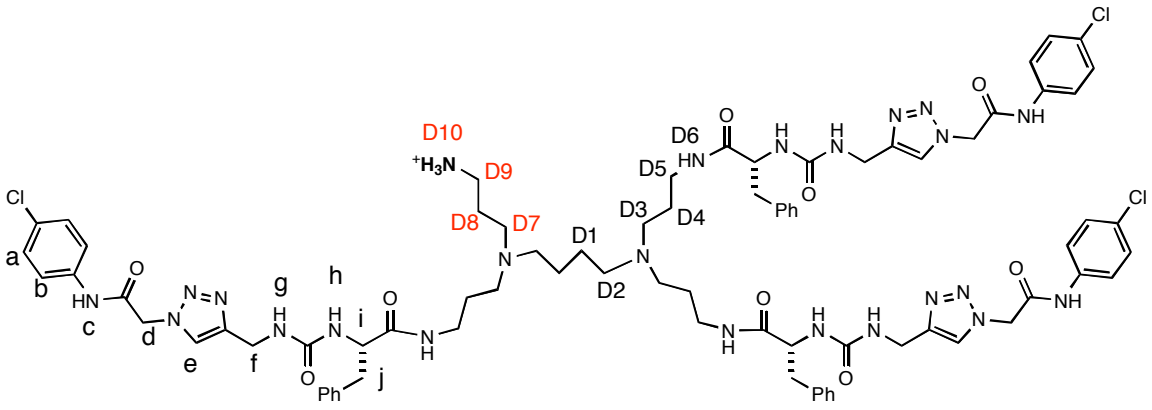
27 **28** **10**

Table 4.6. Purity and Yield Data for Trivalent Molecules 10

compound **10**	alkyne **8**	azide **9**	purity[a] (UV/Sedex, %)	yield[b] (%)
at	a	t	94/96	75
au	a	u	85/97	68
ay	a	y	73/80	59
az	a	z	78/83	52
bt	b	t	88/96	
bu	b	u	90/99	81
by	b	y	94/100	79
bz	b	z	92/99	
bv	b	v	93/100	
bw	b	w	92/100	87
bx	b	x	91/99	
ct	c	t	90/97	
cu	c	u	90/96	
cy	c	y	91/100	80
cz	c	z	85/92	76
cv	c	v	87/95	
cw	c	w	94/100	
cx	c	x	88/95	84

[a]Purity was assessed by HPLC for crude products via monitoring UV absorption at 254 nm and using an evaporative light scattering detector (Sedex). [b]Yield was based on TentaGel S PHB (0.24 mmol/g).

(a)



10cx

(b)

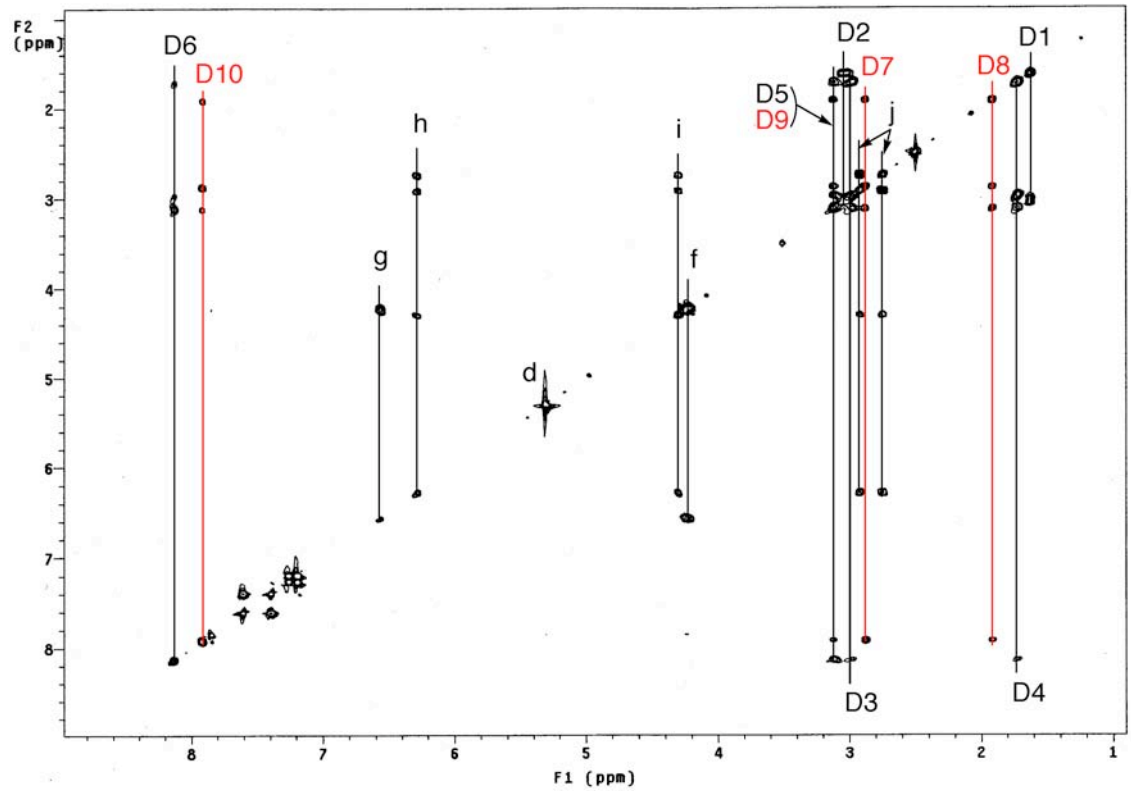


Figure 4.10. Structure (a) and TOCSY spectrum (b) of compound **10cx.**

4.5.4 Immobilization of Multivalent Protein G Mimics

The method used to anchor the peptidomimetics to epoxy-activated sepharose was based on a literature procedure.[100,134] The sepharose resin was first washed several times with pure water to remove some water-soluble additives/impurities. The resin was then mixed with compounds **10** dissolved in 0.2M $NaHCO_3$/DMF (1:1, pH 9.7). The extent of coupling of compounds **10** was monitored in three illustrative cases (**10at**, **10cw** and **10cx**) via RP-HPLC using 1,3-dimethoxybenzene as a standard and following the loss of material from the liquid phase. Data obtained indicated this procedure gave a loading of 3.81 ± 0.02 μmole/g sepharose. Finally, unreacted epoxide groups were capped via treatment with 0.1M Tris (pH 9.4) as described in the literature.[100]

N,*N*-dimethylbenzylamine, 4-methylanisole were also tried as internal standards. They did not give satisfactory results since their peaks by analytical HPLC decrease more than those of compounds **10** as the end of their attachment.

4.6 Summary

An efficient approach for the syntheses of multivalent protein G mimics based on DAB-Am-4 has been developed. The key feature of the synthetic route is that multiligand-functionalized DAB-4 dendrimer with one unfunctionalized amine arm can be obtained by this method. The primary amine of modified DAB dendrimers is very useful for bioassay, such as a fluorescence-attached direct binding assay and sepharose immobilized affinity assay. The reactions are very mild, and only need cheap reagents and gentle shaking. The crude purity is high so that no further purification is needed. This approach is suitable for a parallel synthesis strategy, so it is good to be used for library preparation. A focused library of multiligand-functionalized DAB dendrimers including Amersham's and our designed ligands was prepared. A few compounds were tested by Amersham for their bioactivities. Compound **32** is a promising ligand, which shows a strong binding to the Fab fragment of IgG.

CHAPTER V
SMALL MOLECULE MIMICS OF THE NEUROTROPHINS

5.1 Specific Aims

This project is to resynthesize and synthesize small molecules that mimic the neurotrophins. Many cyclic peptidomimetics were made and sent to Prof. Saragovi for biological assay before I joined Prof. Kevin Burgess's group. Some of them show good activities on Trk receptors in preliminary screening so Prof. Saragovi requires more amounts of these compounds for further biological studies. To provide a probe for a direct binding assay, he also needs fluorescently labeled compounds. Since neurotrophins exist and function as dimers, bivalent turn mimics with a suitable-length linkage are expected to have more selectivity and higher binding affinity to neurotrophin receptors than monomeric molecules. Before making dimers we must first prepare monomers. So the specific goals of this project are to:

a. resynthesize 14 compounds and characterize 12 of them with MS, ^{1}H and ^{13}C NMR, which were previously confirmed by only MS;
b. compare the bioassay data of these compounds made by Hong Boon Lee and me repectively;
c. make five fluorescently labeled compounds;
d. synthesize several monomers exhibiting good binding with Trk receptors for the preparation of the library of dimers.

5.2 Background and Significance

5.2.1 Specificities of Neurotrophins and Their Receptors

The neurotrophins are a family of homologous dimeric growth factors which

include nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3), neurotrophin-4/5 (NT-4/5), neurotrophin-6 and neurotrophin-7.[271-275]

The biological functions of neurotrophins are mediated via binding to two classes of receptors, the p75 neurotrophin receptor ($p75^{NTR}$) and the Trk receptors which are tyrosine kinase receptors.[276] Each neurotrophin binds specifically to different Trk receptors with relative high affinity (K_d: ~10^{-11}M),[277] that is, NGF binds to TrkA, both BDNF and NT-4/5 bind to TrkB while NT-3 binds to TrkC, but can also bind to TrkA and TrkB with lower affinity.[278-280] All neurotrophins also bind to p75NGFR with low affinity constants varying from $K_d = 10^{-9}$M to Kd = 10^{-11}M.[281] The thickness of the arrow line represents the strength of binding between neurotrophins and their receptors in Figure 5.1.

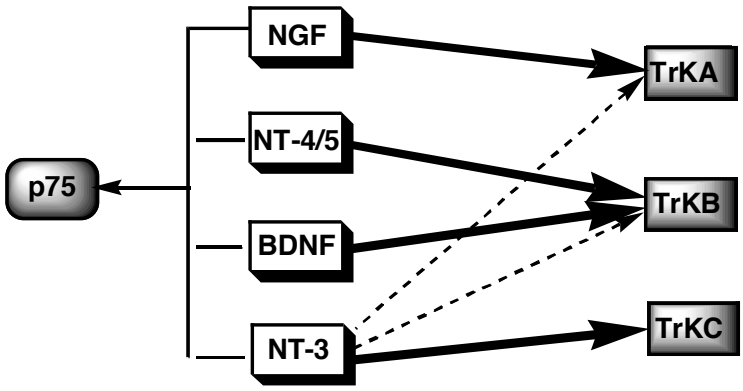


Figure 5.1. Specificities of neurotrophins and their receptors.

5.2.2 The Binding Between Neurotrophins and Their Receptors

Neurotrophins are relatively small polypeptides (~25 kDa and 120 amino acid residues).[272] They share a common structural feature of three intertwined disulfide bridges, exist and function as dimers in the biological system.[272,273] The dimer interface

consists of β-strands in which several hydrophobic residues stabilize the conformation. These hydrophobic residues are highly conserved in neurotrophins.[282] Neurotrophins and their receptors exhibit high sequence homology.[273,276] Crystal structures are available for neurotrophins (NGF dimer: 1btg, NT-3 dimer: 1b8k, NT-4/5 dimer: 1b98) and their receptors (TrkA-d5: 1he7, TrkB-d5: 1wwwb, TrkC-d5: 1wwc) in protein data bank. Binding determinants of neurotrophins to their receptors have been studied by loss-of-function[283-292] and gain-of-function[286,290,293-297] experiments. It has been suggested that the β-turn regions of neurotrophins are highly variable and play a crucial role in the binding affinity and specificities of neurotrophin and receptor interactions because some critical hot spots of neurotrophins are located in these regions. Figure 5.2 shows the loop regions of NGF (pdb: 1www).[298]

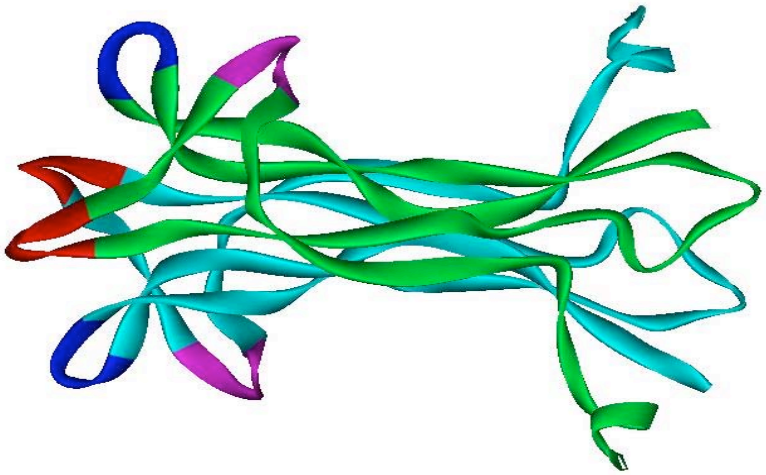


Figure 5.2. NGF with different color highlighted turn regions.

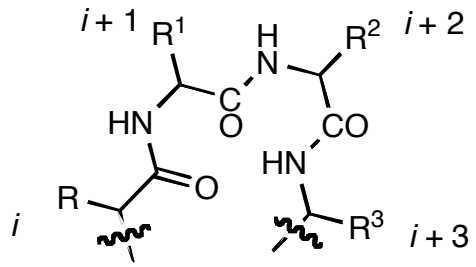
5.2.3 β-Turn Mimics

Neurotrophins develop, regulate and maintain the peripheral and central nervous system. The malfunction of neurotrophins can lead to a number of serious diseases, such as Alzheimer's, stroke,[299-301] neuropathy[302,303] and some forms of

cancer.[304,305] Neurotrophins exhibit a strong ability to promote the survival of peripheral and central neurons during development and to repair the damaged nervous system after neuronal damage, and are regarded as highly potential therapeutic agents in the treatment of nerve injuries and neurodegenerative diseases.[273,306] Nevertheless, since neurotrophins have poor proteolytic stabilities and blood brain barrier permeabilities, they are hard to use as therapeutic agents.[307] In addition to the above drawbacks, they are expensive to produce, and induce undesirable side-effects.[301,308,309] Some of these side-effects come from neurotrophins binding to multiple receptors and activating multiple signaling pathways. So small molecules, which are cheaper and easier to make, and selectively mimic or disrupt the interactions of neurotrophins with their receptors, are of enormous medical and commercial potential. Hot-spot residues are located in the β-turn loops of neurotrophins, so β-turn mimetics are preferable.

The β-turn is one of several important secondary structures, and is often found where a polypeptide chain abruptly reverses direction. It consists of four amino acid residues: *i*, *i*+1, *i*+2 and *i*+3. There are nine types of β-turn according to backbone torsion angle as presented in Table 5.1.[82] Types I and II are the most common β-turn types in proteins.

Table 5.1. Nine Ideal β-Turn Types



β-turn type	backbone torsion angle (°)			
	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
I	-60	-30	-90	0
I'	60	30	90	0
II	-60	120	80	0
II'	60	-120	-80	0
III	-60	-30	-60	-30
III'	60	30	60	30
VIa	-60	120	-90	0
VIb	-120	120	-60	0
VIII	-60	-30	-120	120

To mimic a β-turn, some designs focus on the replacement of the turn hydrogen bond ($_i$C=O---HN$_{i+3}$) with a covalent bond, and keep the side chain topology of *i*+1 and *i*+2 amino acids.[176,310-314] Figure 5.3 shows some examples of these designs. However, most of the compounds are not useful in the exploration of protein-protein interactions because critical side chains corresponding to the target turns are difficult to incorporate to the *i*+1 and *i*+2 amino acids.[315] Moreover, these compounds do not contain an amide-based turn backbone, and most of them were synthesized in solution phase which limits rapid production of libraries.

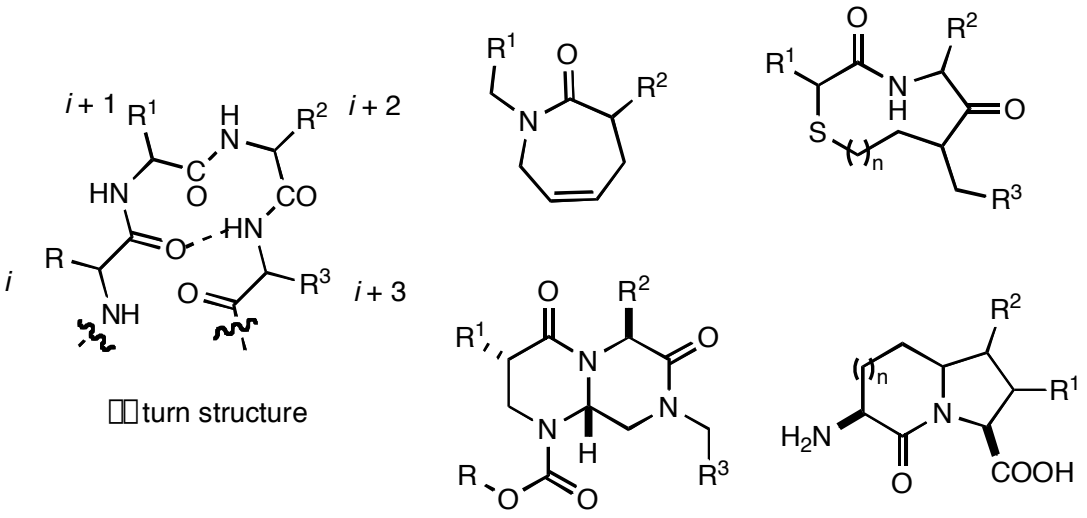


Figure 5.3. Some examples of -turn mimics with a covalent bond instead of turn hydrogen bond.

Some designs retain a turn or partial turn which consists of amino acids in the mimics (Figure 5.4),[316-318] and they were not made via an efficient solid phase strategy. For some compounds, the ring is so flexible that it induces an entropic penalty when the mimics bind to targets. On the other hand, if the compounds are too rigid, this may prevent molecules from adapting to the binding site.

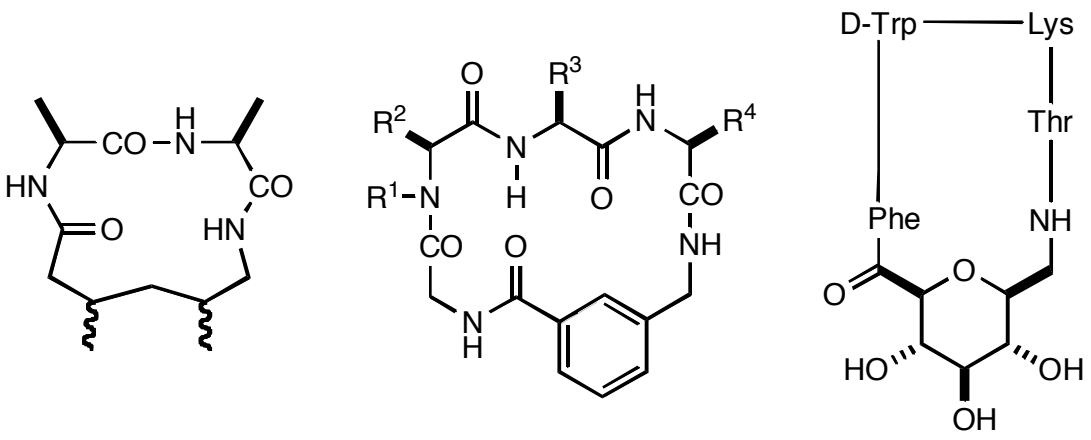


Figure 5.4. Some examples of β-turn mimics retaining a partial turn.

The reasonable solution to this problem is to find compounds which can bind target proteins, then to synthesize and test molecules with suitable amino acid side chains, for example, hot-spots of neurotrophin turn regions, and finally adjust conformational constraints. Studies of conformations of the peptidomimetics and molecular modeling in solution may be useful to guide this process.

Dr. Burgess' design of peptidomimics is composed of two parts: one is peptidic (two or three amino acids), the other is non-peptidic which may improve bioactive conformation. An efficient solid phase route to access a library of β-turn analogs has been developed. The first generation β-turn peptidomimetics with ring-fused C^{10} motifs were made successfully via this method.[171,19,319-326] The results from assay indicate that the first generation β-turn peptidomimetics are able to disrupt some protein-protein interactions. **D3** whose structure was presented in Figure 5.5 is one of the lead compounds.[327] These lead compounds themselves suggest that the peptidic and non-peptidic fragments of the first generation β-turn peptidomimetics could be a very favorable arrangement with regard to protein-protein interactions. Based on the first generation compounds, the second generation β-turn peptidomimetics were designed as

shown in Figure 5.5. The difference is that the first generation compounds are less rigid than the second generation compounds, in which **a**, **b** and **c** bonds are co-planar, and have less degrees of movable freedom. Besides, the nitro group of nitroaryl compounds was replaced with NH_2, CH_3SO_2NH or $NH_2(C{=}NH)NH$, which are normally regarded as conferring drug-likeness to candidates for pharmaceutical development.

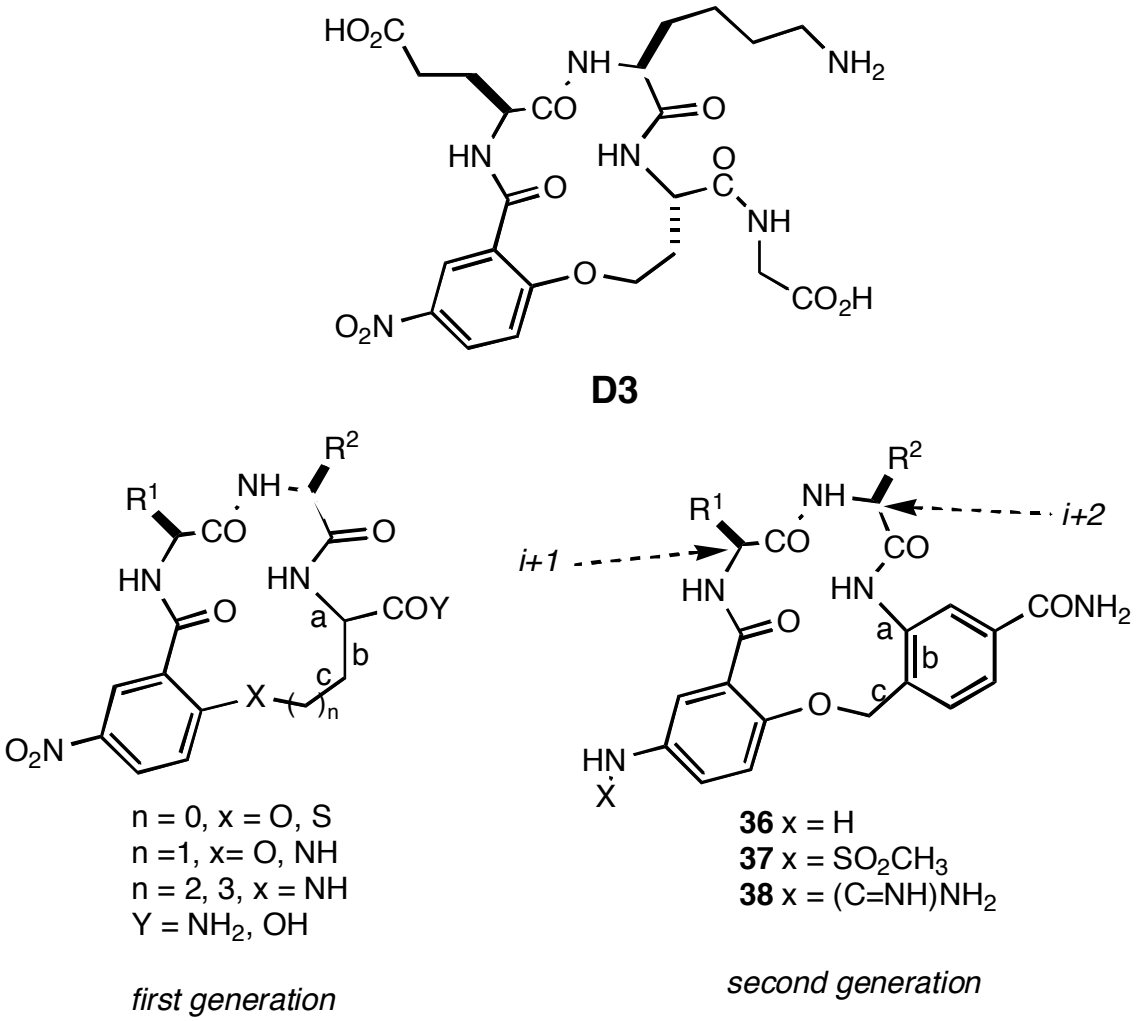


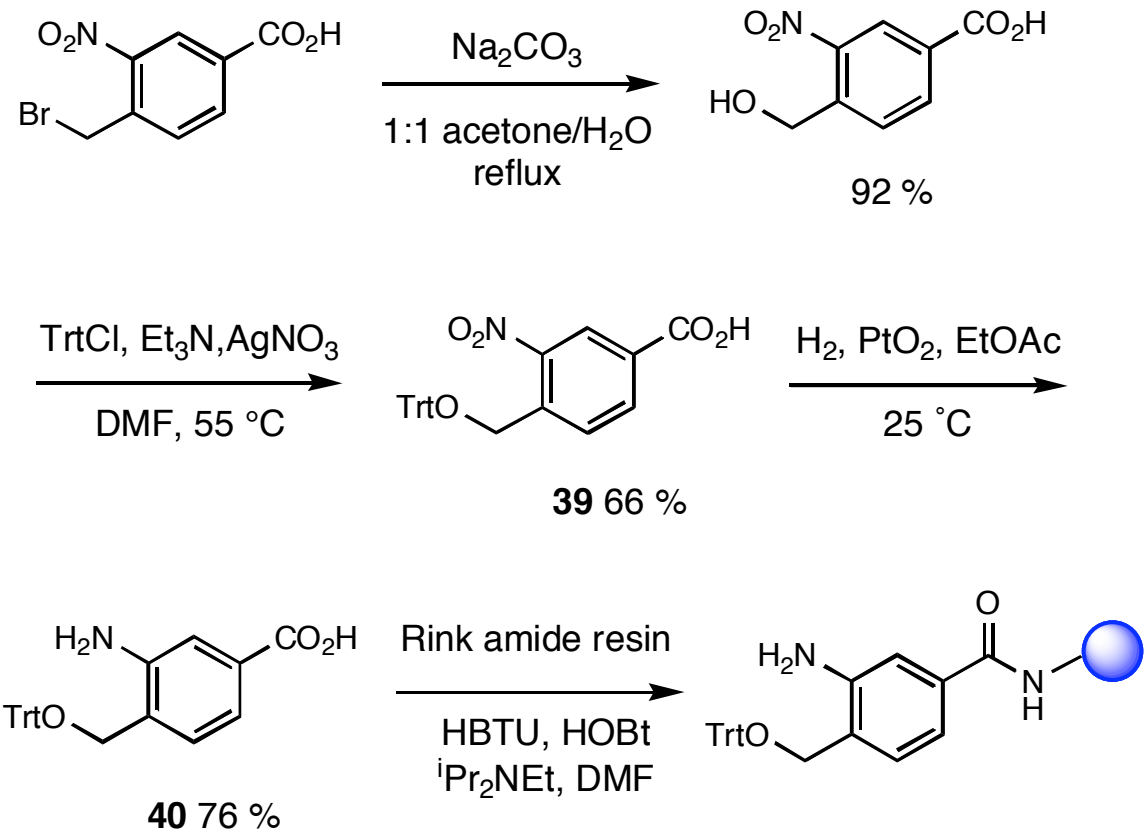
Figure 5.5. Structures of **D3**, first and second generation compounds.

Since Trk receptors are activated by neurotrophins, which are natural dimers, monomer turn analogs of neurotrophins are expected to either exhibit antagonistic effect or have no bioactivity. Therefore, exploring bivalent turn analogs of neurotrophins is more significant in medicinal chemistry. A focused library of dimers, each of which contains two lead monomeric β-turn peptidomimetics was prepared by the Burgess group.[328]

5.3 Synthesis of the Template

Scheme 5.1 describes the syntheses of compound **40** called the template and is the rigid element for all peptidomimetics in this chapter. The template **40** was synthesized via three steps in Trt (triphenylmethyl)-protected form which can be removed using 1% TFA. Hydrolysis[329] of commercially available 4-(bromomethyl)-3-nitrobenzoic acid followed by protection of the resulting alcohol with TrtCl gave compound **39**[330] in 92% and 66% yield respectively. Then hydrogenation of the nitro group of compound **39** using Adam's catalyst (PtO_2) proceeded without significant hydrogenolysis of trityl ether to give template **40** in 76 % yield.[331] Template **40** may be activated with common coupling reagents and coupled to resins functionalized with the Rink handle.[332] This reaction is not complicated by reactions at the free amine group.[333]

Scheme 5.1. Preparation of the Template

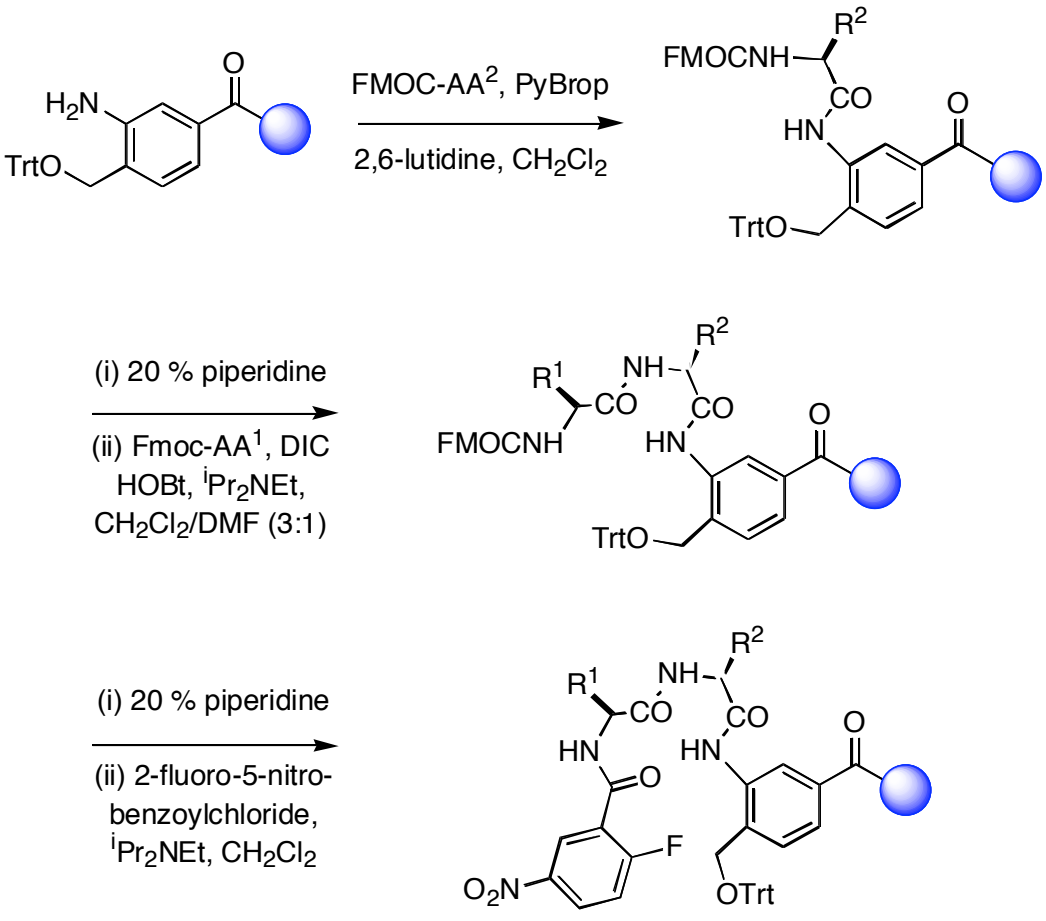


5.4 Solid Phase Syntheses of Peptidomimetics 36-38

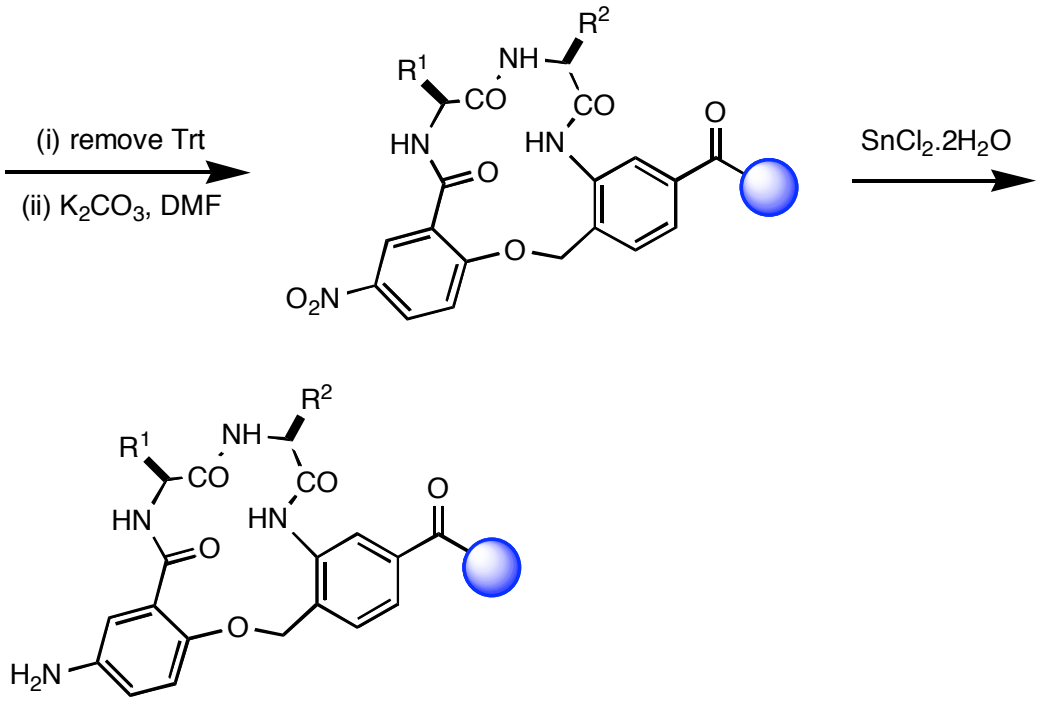
Coupling of the first amino acids to the solid supported template **40** was the most difficult step in the syntheses of these peptidomimetics described in Scheme 5.2. For this step, it was ever found that PyBrop/2,6-lutidine was a superior system;[334] much more excess of 2,6-lutidine was required to prevent loss of the Trt protecting groups during the coupling reaction. After the first coupling reaction was completed, the standard amino acid coupling methods were applied for two more steps, followed by removal of the Trt-protecting group, then base mediated cyclization to give cyclized nitroaryl compounds which was reduced to the corresponding aryl amine compounds on the resin using $SnCl_2 \bullet 2H_2O$/DMF (Scheme 5.2).

The resulting resin supported aryl amines may be cleaved from the resin to give peptidomimetics **36** (Scheme 5. 3). Meanwhile, the side chain protecting groups was also removed in this cleavage using 90 % TFA. These resulting resin supported aryl amines react with methanesulfonyl chloride to give the corresponding sulfonamides **37** (Scheme 5. 3), and react with guanidinylating agents (*N,N'-bis*-BOC-1-guanylpyrazole)[335] to give the guanidines **38** (Scheme 5. 3). The advantages of these reactions are to introduce potential pharmacophores and increase water solubility. Peptidomimetics **36-38** are obviously more water-soluble than nitroaryl compounds.

Scheme 5.2. Preparation of Resin Supported Peptidomimetics 36

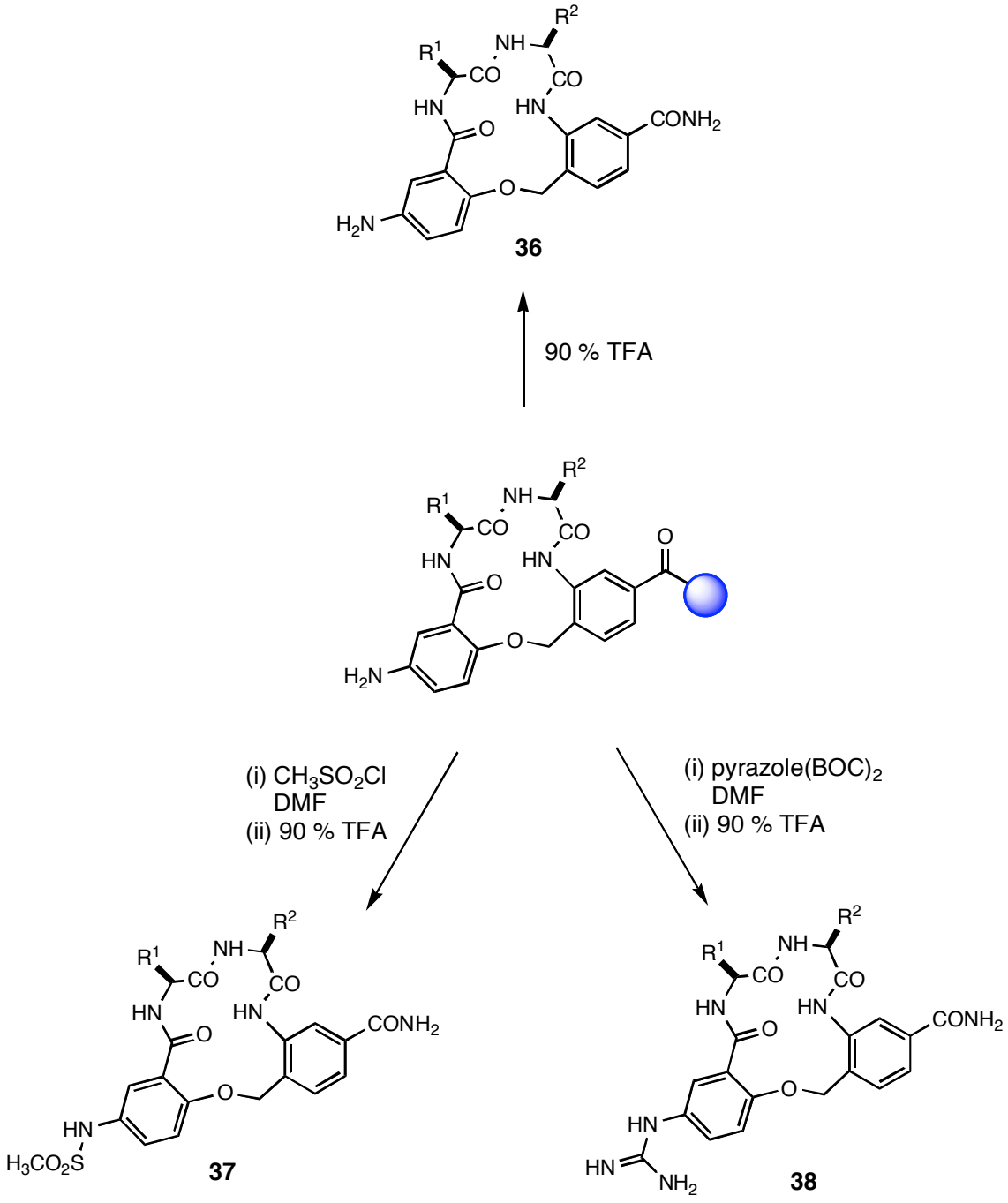


Scheme 5.2. (Continued)



36 with resin

Scheme 5.3. Preparation of Peptidomimetics 36-38



A small library of peptidomimetics **36-38** was reprepared using HypoGel-400-Ram resin via the synthetic strategy depicted in Schemes 5.2 and 5.3. Table 5.2 summarizes yield and purity data for peptidomimetics **36-38**. The purities were generally greater than 60% (UV) for crude materials. After the crudes were purified via preparative HPLC, the purities of most peptidomimetics **36-38** are 100%; and analytical HPLC shows a single peak for most of these compounds. One thing which should be pointed out is that the purities and yields are high when either Lys/Ile or Lys/Thr were used for the dipeptides in the peptidomimetics **36-38**. This work was to only provide samples for bioassay studies, no further chemistry and conformational studies were carried out.

Table 5.2. Summary of Purity and Yield Data for Compounds 36-38[a]

compound	amino acid		purity[b] (%)	purity[c] (%)	isolated yield
	i+1	i+2	(UV)	(UV/Sedex)	(%)
36a	Ile	Lys	90.2	100/100	40
36b	Gly	Lys	61.5	94.6/100	33
36c	Ile	Arg	75.3	91.4/100	23
36d	Arg	Gly	86.3	100/100	45
37a	Ile	Lys	90.2	100/100	73
37b	Ile	Arg	60.7	100/100	13
37c	Lys	Gly	70.6	100/100	28
37d	Arg	Gly	84.1	100/100	38
37e	Lys	Thr	74.7	100/100	39
38a	Ile	Lys	92.4	100/100	50
38b	Gly	Lys	75.6	93.2/94.6	33
38c	Lys	Thr	93.6	100/100	50

[a]HypoGel 400 RAM (0.53 mmol/g) for syntheses of compounds **36-38**. Purity[b] and Purity[c] assessed by HPLC for crude and purified product respectively via monitoring UV absorption at 254 nm and using an evaporative light scattering detector (Sedex).

5.5 Synthesis of Fluoresceinylated Peptidomimetics 41

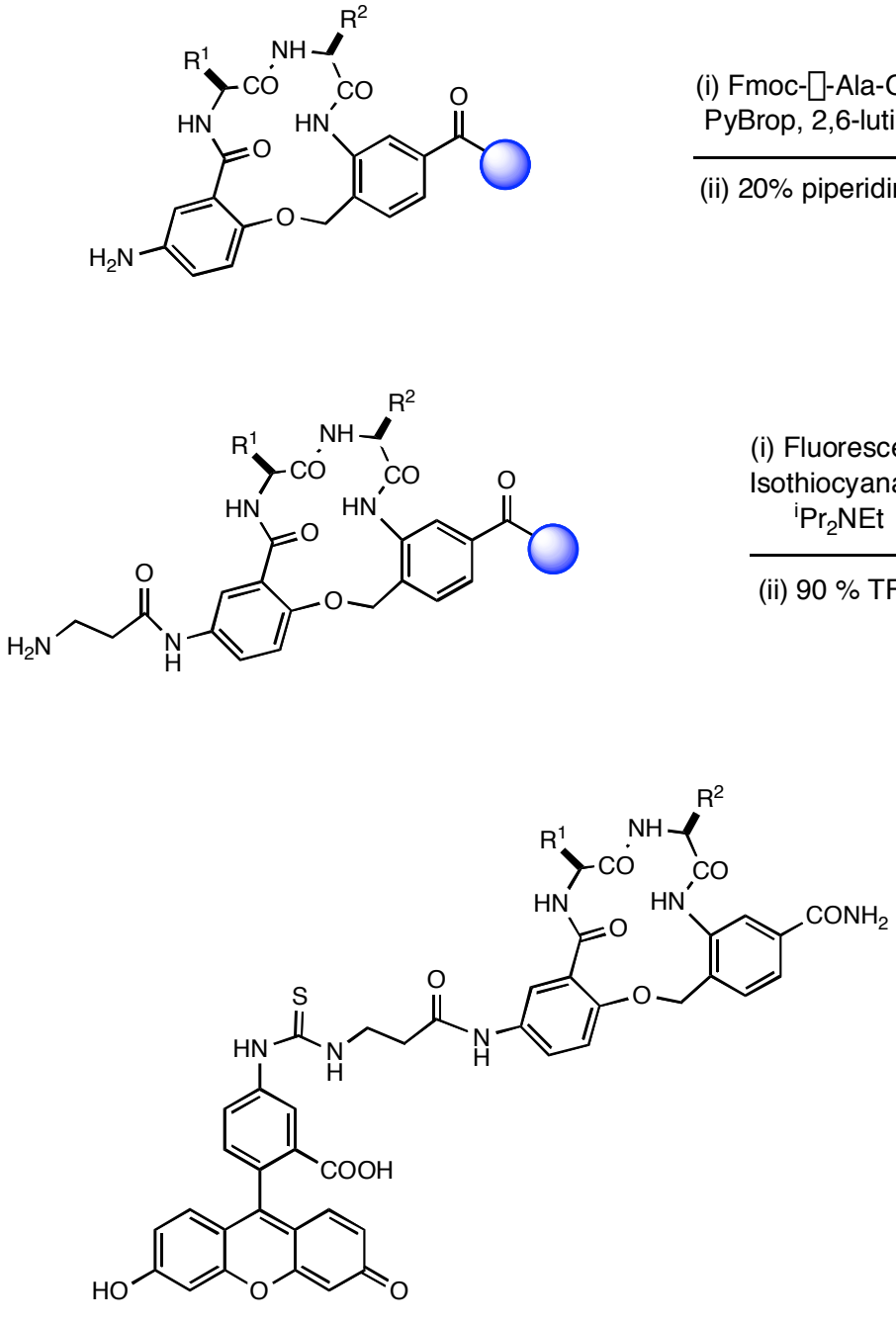
Resin supported aryl amines **36** were first prepared according to the steps depicted in Scheme 5.2. It was then reacted with Fmoc-β Ala-OH using the system (PyBrop/2,6-lutidine), followed by the attachment of fluorescent label using fluorescein isothiocyanate under the base condition. Finally, fluoresceinylated peptidomimetics **41** were released from the resin with simultaneous removal of the side chain protecting groups using 90 % TFA (Scheme 5.4). Application of β-alanine is to avoid the difficulty of the reaction between aryl amine and fluorescein isothiocyanate due to their stereo-size. Purity and yield data were summarized in Table 5.3. Purities of most products are not high even though all products were purified with preparative HPLC. This could be partially attributable to the impurity of fluorescein isothiocyanate.

Table 5.3. Summary of Purity and Yield Data for Compounds 41[a]

Compound	amino acid		purity[b]	isolated yield
41	i+1	i+2	(%)	(%)
a	Ile	Lys	92.4	43
b	Gly	Lys	98.2	31
c	Ile	Arg	89.1	29
d	Arg	Gly	88.3	26
e	Lys	Thr	100	51

[a]HypoGel 400 RAM (0.53 mmol/g) for synthesis of compound **41a-e**. Purity[b] assessed by HPLC for purified product via monitoring UV absorption at 254 nm.

Scheme 5.4. Strategy for Attaching Fluorescent Label to Peptidomimetics



41

5.6 Syntheses of Monomeric Precursors 42 and 43 for Fluoresceinylated Bivalent Turn Mimics of Neurotrophins

Figure 5.6 shows the structures of fluoresceinylated bivalent turn mimics consisting of two monomeric precursors. One of them is an unprotected peptidomimetic with a hand, the other is a fluoresceinylated peptidomimetic with a hand. This research focuses on the large-scale preparation and purification of three monomeric precursors for each. These precursors were then provided to Mookda Pattarawarapan for further assembly of dimers.

Scheme 5.5 depicts preparation of two kinds of monomeric precursors. The resin-supported arylamine **36** (Scheme 5.2) was acylated using Fmoc-Gly-Cl which was introduced to increase the spacer length and flexibility. After deprotection of Fmoc group with 20% piperidine in DMF, Fmoc-Inp-OH was introduced using coupling reagents DIC/HOBt. Deprotection of Fmoc group gave resin supported peptidomimetic derivatives. Each sample in the supported monomeric compound was divided into two portions, one of which was cleaved directly from the resin to give **42**, and the other was treated with dichlorotriazinylaminofluorescein followed by 90% TFA cleavage to give the fluorescently labeled triazine peptidomimetics **43** shown in Scheme 5.5 and Table 5.4.

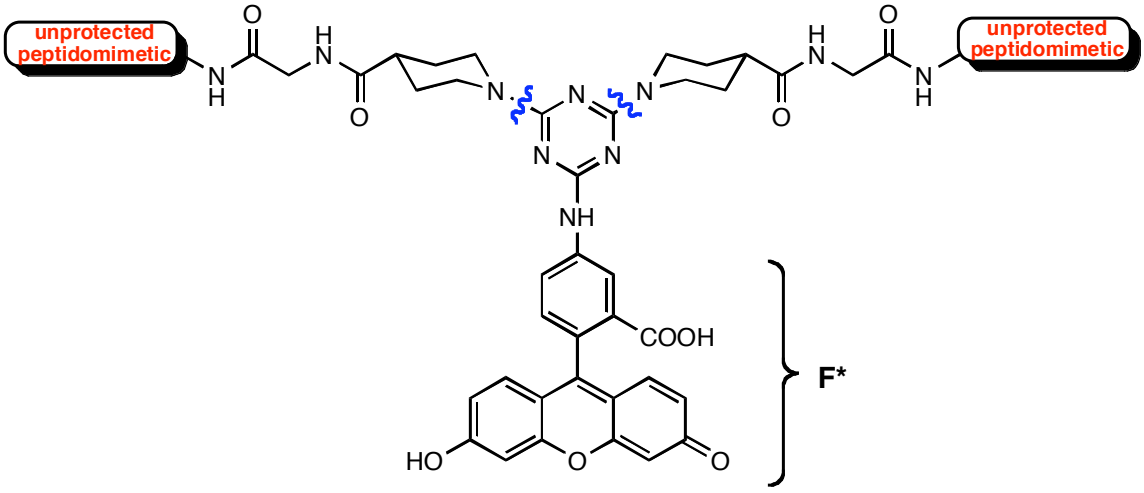
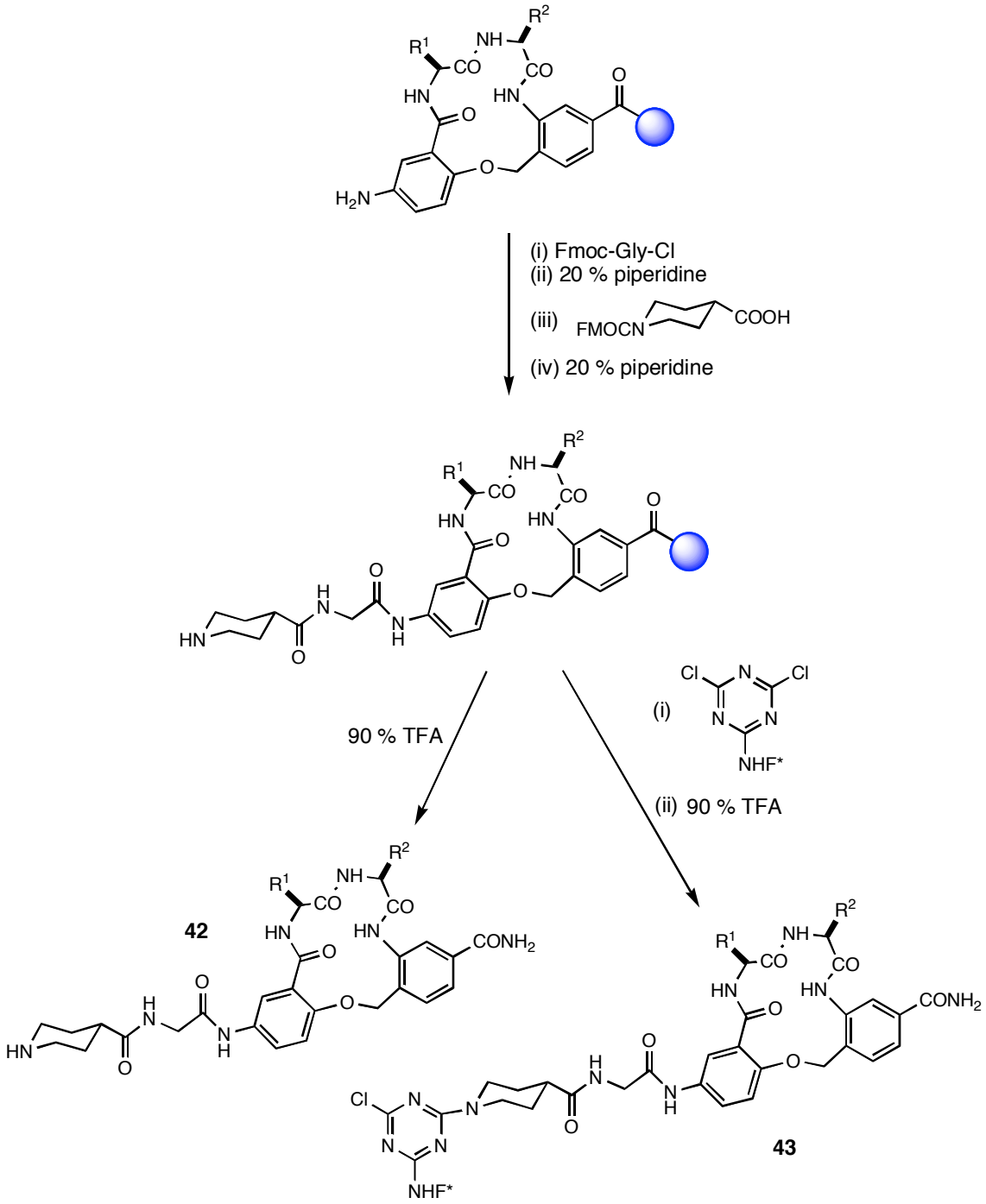


Figure 5.6. Fluoresceinylated bivalent turn mimics.

Table 5.4. Monomers Prepared in Scheme 5.5

compound	amino acid	
42 and **43**	i+1	i+2
a	Ile	Lys
b	Gly	Lys
c	Ile	Arg

Scheme 5.5. Syntheses of Monomeric Precursors 42 and 43



5.7 Comparison of Biological Activities

Biological activity measurements of compounds were carried out by Dr. Uri Saragovi and his co-workers at McGill University in Canada. Cell survival assays[327,336] were selected for these measurements. Cell survival was measured quantitatively by MTT colorimetric assay, optical density readings and calculated relative to 100% survival in 2 nM neurotrophins. Untreated cells in serum free media (SFM), which die without neurotrophins (0% survival), can be rescued by addition of appropriate neurotrophins or mimic ligands with functionality of neurotrophins. Cells were cultured in SFM supplemented with controls (NT-3 and NGF) and mimics with or without additional 0.1nM of appropriate neurotrophins in which NT-3 is for TrKC-expressing cells, and NGF is for TrKA-expresssing cells.

Compounds **36-38** were prepared by Hong Boon Lee and myself respectively. Data from cell survival assays were given in Table 5.5-5.8 and presented graphically in Figure 5.7-5.10. For the same concentration (50 μM) of compounds **36b-d** and **37a**, my data show partial antagonistic activities toward TrkA as they induce the death of the TrkA-expressing E25 cells while Boon's data display partial agonistic activities toward TrkA as they induce the survival of the TrkA expressing cells. When the concentration of compounds **36b-d** was increased to 75 μM, **36b** and **36d** work as agonists toward TrkA, while **36c** is still antagonist as shown by my data. For compounds **36b** and **36d**, both my data and Boon's data indicate partial agonistic activities toward TrkC as the survival of the TrkC-expressing cells was enhanced. In the test of different concentrations (from 0.4 μM to 50 μM) of compounds with 0.1 nM NT-3, compound **36c** shows partial antagonistic activities toward TrkC from my data, but partial agonistic activities from Boon's data.

Table 5.5. Survival Data[a] of E25 TrkA Cells[b] for Compounds 36-38 Made by Different People

compounds	survival of E25 TrkA cell (%)			
	SFM		NGF 0.1nM	
	my data	Boon's data	my data	Boon's data
Untreated	0.00	0.00	0.00	0.00
36a	0.61 ± 1.4	-0.49 ± 0.52	-3.67 ± 0.49	-13.14 ± 6.15
36b	-2.96 ± 0.52 -15.11 (75 μM)	11.15 ± 5.33	-6.45 ± 6.0 37.18 (75 μM)	13.79 ± 8.67
36c	-7.31 ± 2.18 -84 ± 15.04 (75 μM)	6.7 ± 0.09	-11.39 ± 3.25 -62.92 ± 2.67 (75 μM)	2.18 ± 3.18
36d	-2.97 ± 4.15 12.43 ± 8.29 (75 μM)	7.4 ± 0.61	-9.9 ± 0.59 42.35 ± 0.84 (75 μM)	6.73 ± 3.32
37a	-7.34 ± 1.62	5.77 ± 4.94	-14.13 ± 3.99	6.65 ± 6.85
37b	-7.07 ± 3.91	1.91 ± 1.96	-14.48 ± 3.68	-5.70 ± 4.0
37c	-5.8 ± 3.54	0.74 ± 1.36	-4.25 ± 4.24	-0.72 ± 2.72
37d	-6.17 ± 4.52	1.69 ± 1.26	-3.67 ± 4.27	-4.82 ± 2.49
37e	-4.22 ± 2.36	-2.16 ± 3.29	-2.78 ± 2.97	-1.64 ± 3.31
38a	-4.23 ± 3.55	-1.87 ± 2.64	-2.98 ± 5.32	7.34 ± 7.43
38b	-3.22 ± 1.99	-18.47 ± 3.31	-0.51 ± 4.40	-16.96 ± 2.5
38c	-2.34 ± 0.1	-1.35 ± 16.63	-1.78 ± 4.60	-24.1 ± 3.32

[a]Cell survival data obtained for compounds in 50 μM except for the special indication. [b]Cells expressing with the TrkA receptor.

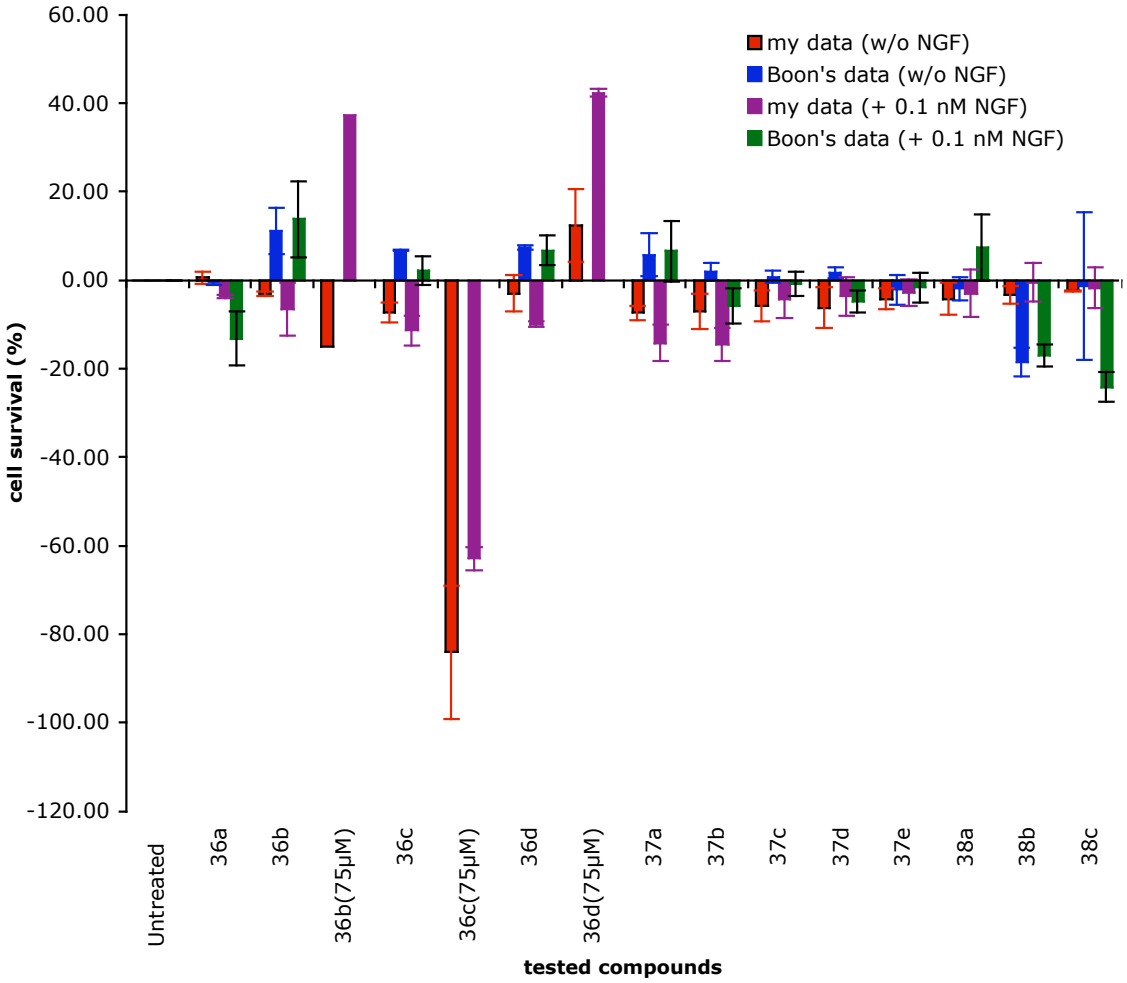


Figure 5.7. Cell survival data for compounds **36**-**38** made by different people in TrkA-expressing cells.

Table 5.6. Survival Data[a] of NIH3T3 TrkC Cells[b] for Compounds 36 and 37 Made by Different People

compounds	survival of NIH3T3 TrkC cell (%)			
	SFM		NT-3 0.1nM	
	my data	Boon's data	my data	Boon's data
Untreated	0.00	0.00	0.00	0.00
36a	0.10 ± 2.13	1.26 ± 0.23	-4.82 ± 1.41	25.75 ± 1.38
36b	1.83 ± 2.05 1.16 ± 7.87 (75 μM)	2.19 ± 3.77	-3.23± 0.59 26.11 ± 3.80 (75 μM)	73.74 ± 24.85
36c	3.26 ± 3.55 -31.77 ± 6.65 (75 μM)	5.74 ± 2.77	-1.43 ± 0.89 -61.23 ± 0.81 (75 μM)	77.65 ± 35.18
36d	-0.18± 4.39 -1.15 ± 3.25 (75 μM)	5.39 ± 3.40	-5.39 ± 2.47 24.09 ± 1.77 (75 μM)	63.28 ± 20.85
37a	2.04 ± 5.28	3.17 ± 4.04	-4.69 ± 8.90	50.04 ± 26.26
37b	-2.22 ± 6.41	1.88 ± 2.11	-6.02 ± 9.39	42.28 ± 33.86

[a]Cell survival data obtained for compounds in 50 μM except for the special indication. [b]Cells expressing with the TrkC receptor.

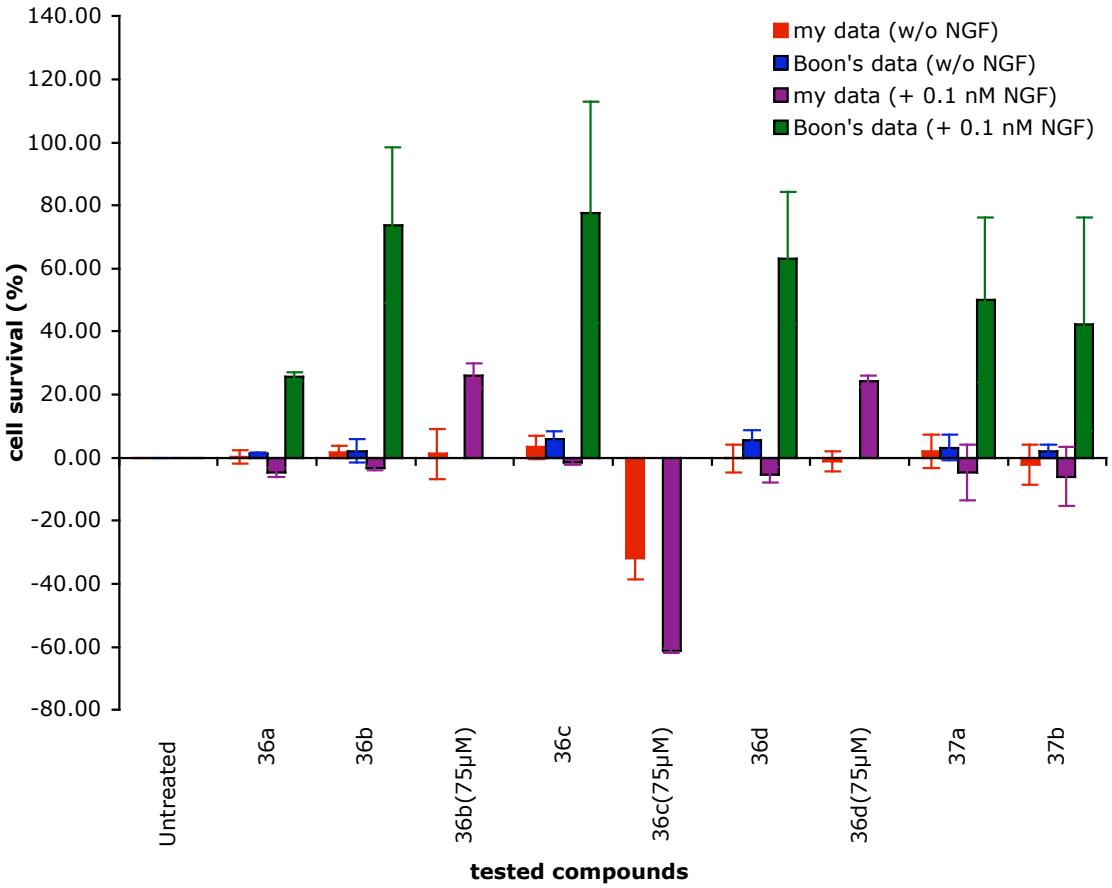


Figure 5.8. Cell survival data for compounds **36** and **37** in TrkC-expressing cells.

Table 5.7. Cell Survival Data for Different Concentrations of Compound 36b Made by Different People

Compound **36b** (μM)	survival of NIH3T3 TrkC cells[a] (%)			
	SFM		NT-3 0.1nM	
	my data	Boon's data	my data	Boon's data
Untreated	0.00	0.00	0.00	0.00
0.4	3.74 ± 4.55	1.01 ± 3.53	-5.28 ± 6.64	5.94 ± 11.51
2	4.92 ± 3.49	2.86 ± 4.73	-0.47± 0.16	12.45 ± 18.37
10	4.49 ± 1.65	4.63 ± 2.46	0.61 ± 1.60	32.22 ± 38.29
50	1.08 ± 0.53	3.94 ± 5.65	-0.81 ± 1.34	39.76 ± 16.11

[a]Cells expressing with the TrkC receptor.

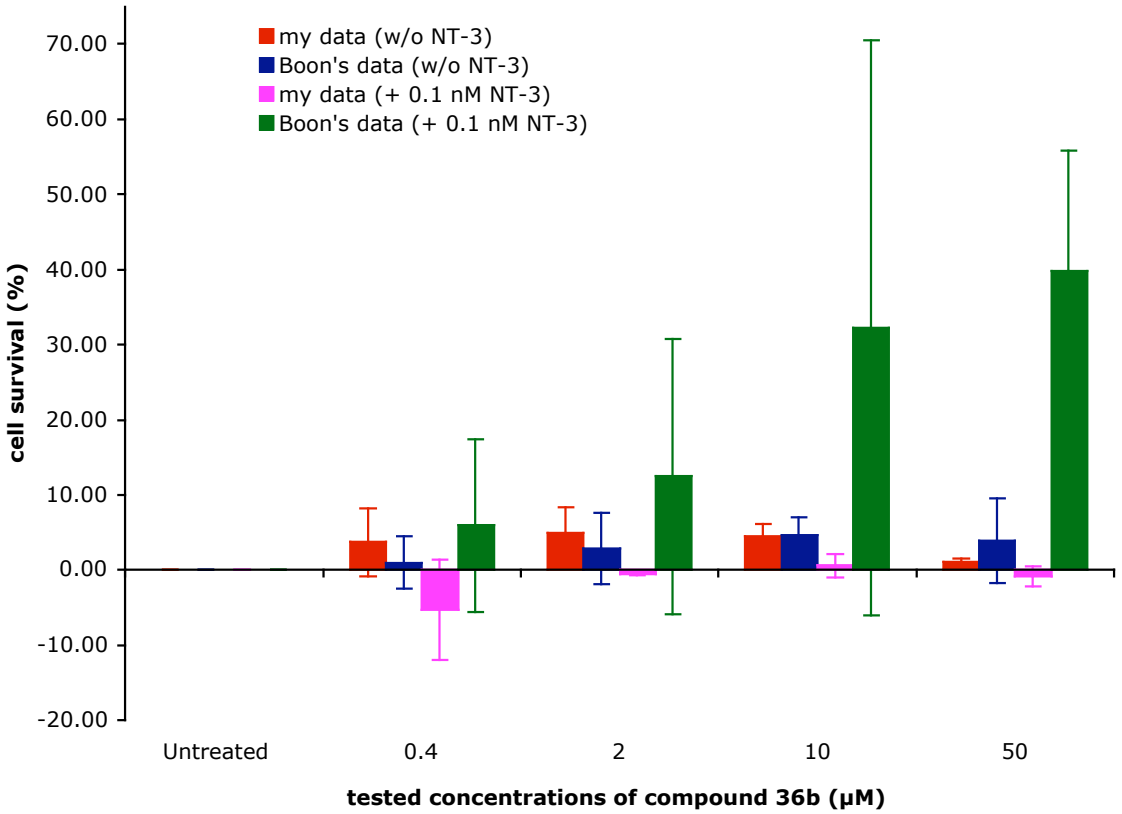


Figure 5.9. Cell survival data for different concentrations of compounds **36b** in TrkC-expressing cells.

Table 5.8. Cell Survival Data for Different Concentrations of Compound 36c Made by Different People

Compound **36c** (μM)	survival of NIH3T3 TrkC cells[a] (%)			
	SFM		NT-3 0.1nM	
	my data	Boon's data	my data	Boon's data
Untreated	0.00	0.00	0.00	0.00
0.4	-1.46 ± 4.12	3.45 ± 2.43	-10.03 ± 13.77	-1.02 ± 6.44
2	0.05 ± 5.26	1.95 ± 2.91	-6.86± 9.09	6.50 ± 13.91
10	-0.16 ± 5.96	4.06 ± 6.66	-6.66 ± 9.62	28.41 ± 35.12
50	1.42 ± 6.20	-0.40 ± 4.33	-11.96 ± 8.30	41.85 ± 33.04

[a]Cells expressing with the TrkC receptor.

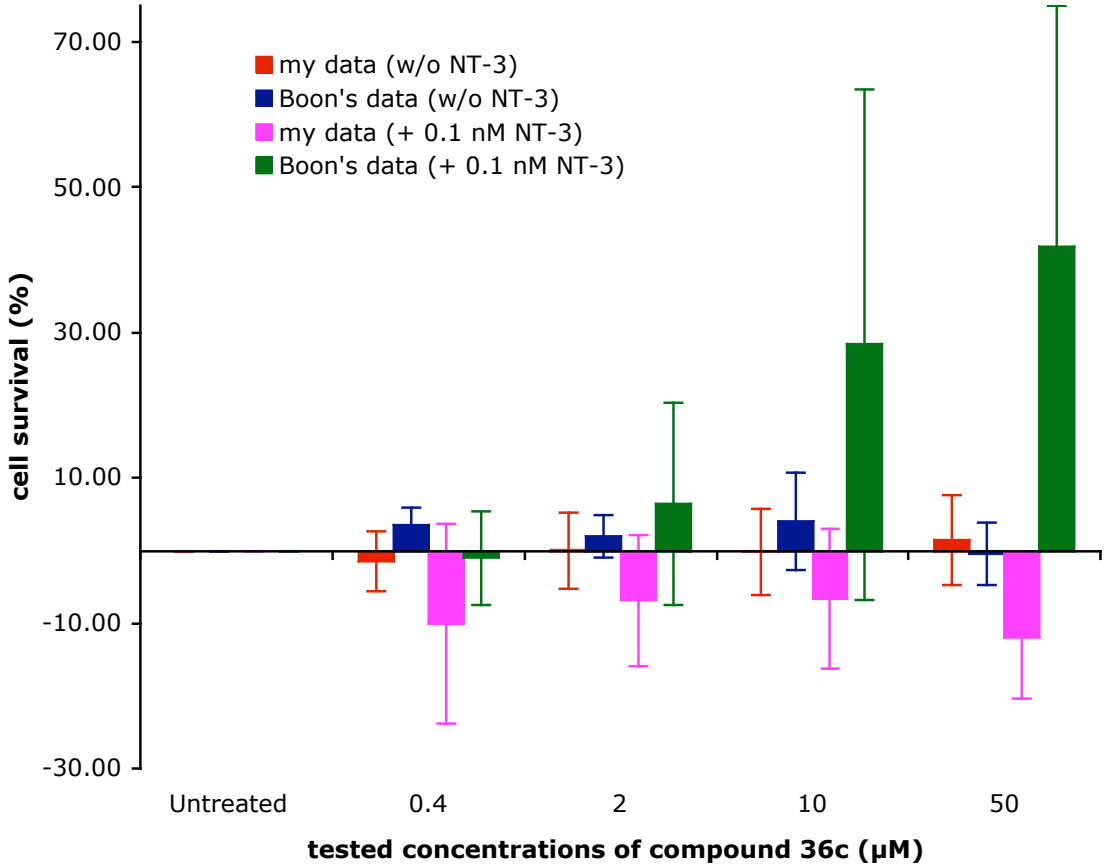


Figure 5.10. Cell survival data for different concentrations of compounds **36c** in TrkC-expressing cells.

5.8 Summary

More than ten compounds which exhibit good bioactivities in the preliminary screening were prepared with appropriate yields. All the compounds were purified by RP-HPLC and characterized with MS and NMR. From the bioassay results provided by Dr. Uri Saragovi, there is no good match between my data and Boon's data. Compounds **36c** and **37a** may be identified as partial antagonist toward both of TrkA and TrkC based on my data, but partial agonistic activities by Boon's data. Both mine and Boon's data show no selectivity for TrkA and TrkC. For compounds **36b** and **36d**, my data indicate that these compounds are partial agonists to both TrkA and TrkC when their

concentration increases to 75 μM, while Boon's data show the same type of activities at the concentration of 50 μM. There are no selective trophic effects between TrkA and TrkC for these compounds. From the concentrations of 0.4 μM to 50 μM of compound **36c**, my data indicate that this compound induces a death of TrkC-expressing NIH3T3 cells exposed to suboptimal NT-3 while Boon's data show a big enhancement of the cell survival. To use the method of fluorescence activated cell sorting (FACScan) assay, five fluoresceinylated peptidomimetics **41** were prepared and purified by RP-HPLC. Three monomeric precursors **42** and **43** were made on large scales for fluoresceinylated bivalent turn mimics of neurotrophins.

CHAPTER VI
CONCLUSIONS

Protein-protein interactions are extremely important in modern drug discovery, synthesis and administration. Hot-spots, which are key residues involved in the interactions, are crucial for mimicking of protein-protein interactions. The approaches in the designs and syntheses of peptidomimetics have been discussed. Generally, the following strategy were adopted in each of my projects: identify protein-protein interactions and hot-spots, estimate the geometry of amino acids in the hot-spots, design mimics based on hot-spots and molecular modeling, synthesize a small library of mimics, analyze mimics including conformational study and modeling, and screen the biological activities of mimics.

Computer-aided design of mimics for protein-protein interactions is an obvious opportunity in modern medicinal chemistry. The QMD technique, which effectively explores the conformations of molecules and has been used for many years in Dr Burgess' group based on Quanta/CHARMm package, can now be used on the Insight II platform. A variety of force fields, such as CVFF, CFF and CHARMm, are available for QMD calculation. The test data for the compounds reported in this dissertation have demonstrated that the exploration of QMD technique in Insight II platform is successful.

The interaction between expensive protein A and IgG is a typical protein-protein interaction that has been used to purify IgG for therapeutic application. Seven peptidomimetics were designed to mimic the helix-loop-helix region of protein A; all seven were synthesized with high yields on a solid support. The key step containing the coupling of FMOC-Aib-OH works well using TFFH. The conformations of four peptidomimetics were analyzed by CD, NMR spectroscopy and molecular modeling. The results of CD, NMR and QMD studies are consistent with one another. As discussed in Chapter III, these data indicate that these peptidomimetics have partial α-helix secondary structures, and the addition of hexafluoroisopropanol to the buffered

water sample has proven to stabilize the α-helix of peptidomimetics. Excitingly, the shortest peptidomimetic **6** was the most helical.

Protein G is another expensive natural ligand for current IgG purification. Since it binds IgG in a completely different way from that of protein A, *ie* a huge binding area, several H-bonds, and both helix and β-sheet of protein G/domain III containing hot-spots, the design of ligands to mimic these discontinuous hot-spots is a great challenge. Poly(propylene imine) dendrimers were selected as scaffolds. Monovalent ligands based on affinity modeling and Amersham's compounds were designed. A new methodology for the formation of multivalent small molecule protein G mimics on a solid support has been developed. As discussed in Chapter IV, the concentration of poly(propylene imine) dendrimers proved to be a key issue. Only click chemistry was used during the attachment of build blocks. The mild reaction conditions, high purity and high yield demonstrated that this methodology is efficient and very useful for the preparation of a library of compounds. However, this approach does not work well for high generations of poly(propylene imine) dendrimers as scaffolds. A few protein G multivalent mimics with Amersham's designed ligands and my peptidomimetic **6** discussed in Chapter III were tested for binding to the Fab fragment of IgG. The strong binding result indicates that peptidomimetic **6** functionalized trivalent **32** (in Chapter IV) is a very promising ligand for IgG purification.

The interactions between neurotrophins and their receptors are another example of protein-protein interactions studied in our lab. The β-turn regions of neurotrophins were believed to be hot-spots in binding to the receptors. Twelve compounds were resynthesized and characterized. In order to rapidly identify compounds and check for their affinity to our protein targets, five fluoresceinylated peptidomimetics were prepared. The biological data obtained for the compounds which were synthesized by Hong Boon Lee and resynthesized by me, respectively, were compared. Generally, there is a poor match between Boon's data and my data. If Boon's data support partially agonistic activities of a compound to either TrkA or C, my data support partially antagonistic activities. For the requirement of the library of bivalent turn mimics, three

lead monomeric precursors **42** and fluorescently labeled compounds **43** were prepared. The dimers display enhanced binding abilities compared to monovalent ligands.[328]

REFERENCES

(1) Eyster, K. M. *Biochem. Pharmacol.* **1998**, *55*, 1927-1938.

(2) Klemm, J. D.; Schreiber, S. L.; Crabtree, G. R. *Annu. Rev. Immunol.* **1998**, *16*, 569-592.

(3) Souroujon, M. C.; Mochly-Rosen, D. *Nat. Biotechnol.* **1998**, *16*, 919-924.

(4) Dall'Acqua, W.; Goldman, E. R.; Lin, W.; Teng, C.; Tsuchiya, D.; Li, H.; Ysern, X.; Braden, B. C.; Li, Y.; Smith-Gill, S. J.; Mariuzza, R. A. *Biochemistry* **1998**, *37*, 7981-7991.

(5) Salzmann, M.; Bachmann, M. F. *Mol. Immunol.* **1998**, *35*, 271-277.

(6) Jones, S.; Thornton, J. M. *Prog. Biophys. Mol. Biol.* **1995**, *63*, 31-65.

(7) Berg, T. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 2462-2481.

(8) Cohen, F. E.; Prusiner, S. B. *Annu. Rev. Biochem.* **1998**, *67*, 793-819.

(9) Loregian, A.; Mardsen, H. S.; Palu, G. *Rev. Med. Virol.* **2002**, *12*, 239-262.

(10) Cochran, A. G. *Chem. Biol.* **2000**, *7*, R85-94.

(11) Bhalla, U. S.; Iyengar, R. *Science* **1999**, *283*, 381-387.

(12) Stites, W. E. *Chem. Rev.* **1997**, *97*, 1233-1250.

(13) Gadek, T. R.; Nicholas, J. B. *Biochem. Pharmacol.* **2003**, *65*, 1-8.

(14) Preissner, R.; Goede, A.; Frommel, C. *J. Mol. Biol.* **1998**, *280*, 535-550.

(15) Kaul, D. R.; Cinti, S. K.; Carver, P. L.; Kazanjian, P. H. *Pharmacotherapy* **1999**, *19*, 281-298.

(16) Toogood, P. L. *J. Med. Chem.* **2002**, *45*, 1543-1558.

(17) Cochran, A. G. *Curr. Opin. Chem. Biol.* **2001**, *5*, 654-659.

(18) Zutshi, R.; Brickner, M.; Chmielewski, J. *Curr. Opin. Chem. Biol.* **1998**, *2*, 62-66.

(19) Feng, Y.; Burgess, K. *Chem. Eur. J.* **1999**, *5*, 3261-3272.

(20) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13-20.

(21) McCoy, A. J.; Chandana Epa, V.; Colman, P. M. *J. Mol. Biol.* **1997**, *268*, 570-584.

(22) Lawrence, M. C.; Colman, P. M. *J. Mol. Biol.* **1993**, *234*, 946-950.

(23) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155-164.

(24) Lo Conte, L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177-2198.

(25) Chothia, C.; Janin, J. *Nature* **1975**, *256*, 705-708.

(26) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1-9.

(27) Clackson, T.; Wells, J. A. *Science* **1995**, *267*, 383-386.

(28) Massova, I.; Kollman, P. A. *J. Am. Chem Soc.* **1999**, *121*, 8133-8143.

(29) Vaughan, C. K.; Buckle, A. M.; Fersht, A. R. *J. Mol. Biol.* **1999**, *286*, 1487-1506.

(30) Ortiz, M. A.; Light, J.; Maki, R. A.; Assa-Munt, N. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2740-2745.

(31) Ivanov, Y. D.; Kanaeva, I. P.; Karuzina, II; Archakov, A. I.; Hoa, G. H.; Sligar, S. G. *Arch. Biochem. Biophys.* **2001**, *391*, 255-264.

(32) Leavitt, S.; Freire, E. *Curr. Opin. Struct. Biol.* **2001**, *11*, 560-566.

(33) Sloan, D. J.; Hellinga, H. W. *Protein Eng.* **1998**, *11*, 819-823.

(34) Fancy, D. A.; Kodadek, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6020-6024.

(35) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. *Science* **1989**, *245*, 616-621.

(36) Todd, M. J.; Semo, N.; Freire, E. *J. Mol. Biol.* **1998**, *283*, 475-488.

(37) Shultz, M. D.; Chmielewski, J. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2431-2436.

(38) Zhang, Z. Y.; Poorman, R. A.; Maggiora, L. L.; Heinrikson, R. L.; Kezdy, F. J. *J. Biol. Chem.* **1991**, *266*, 15591-15594.

(39) Shultz, M. D.; Bowman, M. J.; Ham, Y. W.; Zhao, X.; Tora, G.; Chmielewski, J. *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 2710-2713.

(40) Darling, A. J.; McKay, E. M.; Ingemarson, R.; Preston, V. G. *Virus Genes* **1989**, *2*, 187-194.

(41) Filatov, D.; Ingemarson, R.; Graslund, A.; Thelander, L. *J. Biol. Chem.* **1992**, *267*, 15816-15822.

(42) Moss, N.; Deziel, R.; Adams, J.; Aubry, N.; Bailey, M.; Baillet, M.; Beaulieu, P.; DiMaio, J.; Duceppe, J. S.; Ferland, J. M.; et al. *J. Med. Chem.* **1993**, *36*, 3005-3009.

(43) Moss, N.; Beaulieu, P.; Duceppe, J. S.; Ferland, J. M.; Gauthier, J.; Ghiro, E.; Goulet, S.; Grenier, L.; Llinas-Brunet, M.; Plante, R.; et al. *J. Med. Chem.* **1995**, *38*, 3617-3623.

(44) Moss, N.; Beaulieu, P.; Duceppe, J. S.; Ferland, J. M.; Garneau, M.; Gauthier, J.; Ghiro, E.; Goulet, S.; Guse, I.; Jaramillo, J.; Llinas-Brunet, M.; Malenfant, E.; Plante, R.; Poirier, M.; Soucy, F.; Wernic, D.; Yoakim, C.; Deziel, R. *J. Med. Chem.* **1996**, *39*, 4173-4180.

(45) Liuzzi, M.; Deziel, R.; Moss, N.; Beaulieu, P.; Bonneau, A. M.; Bousquet, C.; Chafouleas, J. G.; Garneau, M.; Jaramillo, J.; Krogsrud, R. L.; et al. *Nature* **1994**, *372*, 695-698.

(46) Fisher, K. L.; Lu, J.; Riddle, L.; Kim, K. J.; Presta, L. G.; Bodary, S. C. *Mol. Biol. Cell* **1997**, *8*, 501-515.

(47) Casasnovas, J. M.; Stehle, T.; Liu, J. H.; Wang, J. H.; Springer, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4134-4139.

(48) McMurray, R. W. *Semin Arthritis Rheum* **1996**, *25*, 215-233.

(49) Oppenheimer-Marks, N.; Lipsky, P. E. *Clin. Immunol. Immunopathol.* **1996**, *79*, 203-210.

(50) Adler, M.; Carter, P.; Lazarus, R. A.; Wagner, G. *Biochemistry* **1993**, *32*, 282-289.

(51) Dennis, M. S.; Carter, P.; Lazarus, R. A. *Proteins* **1993**, *15*, 312-321.

(52) Burdick, D. J. WO 9949856, **1999**, p 9-16.

(53) Gadek, T. R.; Burdick, D. J.; McDowell, R. S.; Stanley, M. S.; Marsters, J. C., Jr.; Paris, K. J.; Oare, D. A.; Reynolds, M. E.; Ladner, C.; Zioncheck, K. A.; Lee, W. P.; Gribling, P.; Dennis, M. S.; Skelton, N. J.; Tumas, D. B.; Clark, K. R.; Keating, S. M.; Beresini, M. H.; Tilley, J. W.; Presta, L. G.; Bodary, S. C. *Science* **2002**, *295*, 1086-1089.

(54) Gross, A.; McDonnell, J. M.; Korsmeyer, S. J. *Genes Dev*. **1999**, *13*, 1899-1911.

(55) Reed, J. C. *Trends Mol. Med*. **2001**, *7*, 314-319.

(56) Fesik, S. W. *Cell* **2000**, *103*, 273-282.

(57) Muchmore, S. W.; Sattler, M.; Liang, H.; Meadows, R. P.; Harlan, J. E.; Yoon, H. S.; Nettesheim, D.; Chang, B. S.; Thompson, C. B.; Wong, S. L.; Ng, S. L.; Fesik, S. W. *Nature* **1996**, *381*, 335-341.

(58) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; Thompson, C. B.; Fesik, S. W. *Science* **1997**, *275*, 983-986.

(59) Wang, J. L.; Liu, D.; Zhang, Z. J.; Shan, S.; Han, X.; Srinivasula, S. M.; Croce, C. M.; Alnemri, E. S.; Huang, Z. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7124-7129.

(60) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. *J. Am. Chem. Soc*. **2002**, *124*, 11838-11839.

(61) Hendrickson, J. B. *J. Am. Chem. Soc* **1961**, *83*, 4537-4547.

(62) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekaran, V. *J. Mol. Biol*. **1963**, *7*, 95-99.

(63) De Santis, P.; Giglio, E.; Liquori, A. M.; Ripamonte, A. *J. Polym. Sci. Part A* **1963**, *1*, 1383-1404.

(64) Williams, D. E. *Chem. Phys*. **1981**, *55*, 163-167.

(65) Jorgensen, W. L. *J. Am. Chem. Soc*. **1981**, *103*, 335-340.

(66) Olson, W. K. *J. Am. Chem. Soc*. **1982**, *104*, 278-286.

(67) Van Gunsteren, W. F.; Karplus, M. *Biochemistry* **1982**, *21*, 2259-2274.

(68) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117-135.

(69) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493-2499.

(70) Florián, J.; Leszczynski, J. *J. Biomol. Struct. Dyn.* **1995**, *12*, 1055-1062.

(71) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Comput. Chem.* **1996**, *17*, 841-850.

(72) Florián, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010-3017.

(73) Sponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161-3164.

(74) Nikiforovich, G. V. *Int. J. Pept. Protein Res.* **1994**, *44*, 513-531.

(75) Brooks, B.; Bruccoleri, R.; Olafson, B. S., D.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187-217.

(76) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1913-1937.

(77) Smith, P. E.; Dang, L. X.; Pattitt, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 67-73.

(78) Verlet, L. *Phys. Rev.* **1967**, *159*, 98-103.

(79) O'Connor, S. D.; Smith, P. E.; Al-Obeidi, F.; Pettitt, B. M. *J. Med. Chem.* **1992**, *35*, 2870-2881.

(80) http://www.accelrys.com/doc, "Forcefield-Based Simulations".

(81) Chou, P. Y.; Fasman, G. D. *J. Mol. Biol.* **1997**, *115*, 135-175.

(82) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221-232.

(83) Richardon, J. S. *Adv. Protein Chem.* **1981**, *34*, 167-339.

(84) Lin, C. C.; Wu, J. C.; Chang, T. T.; Chang, W. Y.; Yu, M. L.; Tam, A. W.; Wang, S. C.; Huang, Y. H.; Chang, F. Y.; Lee, S. D. *J. Clin. Microbiol.* **2000**, *38*, 3915-3918.

(85) Lu, B.; Smyth, M. R.; O'Kennedy, R. *Analyst* **1996**, *121*, 29R-32R.

(86) Yarmush, M. L.; Weiss, A. M.; Antonsen, K. P.; Odde, D. J.; Yarmush, D. M. *Biotechnol. Adv.* **1992**, *10*, 413-446.

(87) Askari, M.; Alarie, J. P.; Moreno-Bondi, M.; Vo-Dinh, T. *Biotechnol. Prog.* **2001**, *17*, 543-552.

(88) Stokes, D. L.; Griffin, G. D.; Vo-Dinh, T. *Fresenius J. Anal. Chem.* **2001**, *369*, 295-301.

(89) Mendoza, L. G.; McQuary, P.; Mongan, A.; Gangadharan, R.; Brignac, S.; Eggers, M. *Biotechniques* **1999**, *27*, 778-788.

(90) Banerjee, A.; Srilatha, N. S.; Murthy, G. S. *Biochim. Biophys. Acta.* **2002**, *1569*, 21-30.

(91) van Dijk, M. A.; van de Winkel, J. G. *Curr. Opin. Chem. Biol.* **2001**, *5*, 368-374.

(92) Cragg, M. S.; French, R. R.; Glennie, M. J. *Curr. Opin. Immunol.* **1999**, *11*, 541-547.

(93) Vaswani, S. K.; Hamilton, R. G. *Ann. Allergy Asthma Immunol.* **1998**, *81*, 105-119.

(94) Kohler, G.; Milstein, C. *Nature* **1975**, *256*, 495-497.

(95) Guddat, L. W.; Herron, J. N.; Edmundson, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4271-4275.

(96) Bond, A.; Jones, M. G.; Hay, F. C. *J. Immunol. Methods* **1993**, *166*, 27-33.

(97) Boden, V.; Winzerling, J. J.; Vijayalakshmi, M.; Porath, J. *J. Immunol. Methods* **1995**, *181*, 225-232.

(98) Labrou, N.; Clonis, Y. D. *J. Biotechnol.* **1994**, *36*, 95-119.

(99) Huse, K.; Bohme, H. J.; Scholz, G. H. *J. Biochem. Biophys. Methods* **2002**, *51*, 217-231.

(100) Verdoliva, A.; Cassani, G.; Fassina, G. *J. Chromatogr. B Biomed. Appl.* **1995**, *664*, 175-183.

(101) Labrou, N. E. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2003**, *790*, 67-78.

(102) Langone, J. J. *Adv. Immunol.* **1982**, *32*, 157-252.

(103) Kihira, Y.; Aiba, S. *J. Chromatogr.* **1992**, *597*, 277-283.

(104) Deisenhofer, J. *Biochemistry* **1981**, *20*, 2361-2370.

(105) Saha, K.; Bender, F.; Gizeli, E. *Anal. Chem.* **2003**, *75*, 835-842.

(106) Langone, J. J. *J. Immunol. Methods* **1982**, *55*, 277-296.

(107) Hale, G.; Drumm, A.; Harrison, P.; Phillips, J. *J. Immunol. Methods* **1994**, *171*, 15-21.

(108) Sjodahl, J. *Eur. J. Biochem.* **1977**, *73*, 343-351.

(109) Moks, T.; Abrahmsen, L.; Nilsson, B.; Hellman, U.; Sjoquist, J.; Uhlen, M. *Eur. J. Biochem.* **1986**, *156*, 637-643.

(110) Deisenhofer, J.; Jones, T. A.; Huber, R.; Sjodahl, J.; Sjoquist, J. *Hoppe-Seyler's Z. Physiol. Chem.* **1978**, *359*, 975-985.

(111) Torigoe, H.; Shimada, I.; Saito, A.; Ssato, M.; Arata, Y. *Biochemistry* **1990**, *29*, 8787-8793.

(112) Doolittle, R. F. *Protein Sci.* **1992**, *1*, 191-200.

(113) Myhre, E. B. *Bacterial Immunoglobulin Binding Proteins*; Boyle, M. D. P., Ed.; Academic Press: London, 1990; Vol. 1, p 243-56.

(114) Ljungberg, U. K.; Jansson, B.; Niss, U.; Nilsson, R.; Sandberg, B. E.; Nilsson, B. *Mol. Immunol.* **1993**, *30*, 1279-1285.

(115) Sasso, E. H.; Silvermann, G. J.; Mannik, M. *J. Immunol.* **1991**, *147*, 1877-1883.

(116) Harboe, M.; Folling, I. *Scand. J. Immunol.* **1974**, *3*, 471-482.

(117) Gouda, H.; Shiraishi, M.; Takahashi, H.; Kato, K.; Torigoe, H.; Arata, Y.; Shimada, I. *Biochemistry* **1998**, *37*, 129-136.

(118) Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arata, Y.; Shimada, I. *Biochemistry* **1992**, *31*, 9665-9672.

(119) Cedergren, L.; Andersson, R.; Jansson, B.; Uhlen, M.; Nilsson, B. *Protein Eng*. **1993**, *6*, 441-448.

(120) Li, R.; Dowd, V.; Stewart, D. J.; Burton, S. J.; Lowe, C. R. *Nat. Biotechnol*. **1998**, *16*, 190-195.

(121) Fuglistaller, P. *J. Immunol. Methods* **1989**, *124*, 171-177.

(122) Godfrey, M. A. J.; Kwasowski, P.; Clift, R.; Marks, V. *Immunol. Methods* **1992**, *149*, 21-27.

(123) Porath, J.; Belew, M. *Trends in Biotechnol*. **1987**, *5*, 225-229.

(124) El-kak, A.; Vijayalakshmi, M. A. *J. Chromatogr, Biomed. Appl*. **1991**, *570*, 29-41.

(125) Ngo, T. T.; Khatter, N. *Appl. Biochem. Biotechnol*. **1991**, *30*, 111-119.

(126) Ngo, T. T.; Khatter, N. *J. Chromatogr*. **1992**, *597*, 101-109.

(127) Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84-86.

(128) Fassina, G.; Verdoliva, A.; Odierna, M. R.; Ruvo, M.; Cassini, G. *J. Mol. Recognit*. **1996**, *9*, 564-569.

(129) Palombo, G.; De Falco, S.; Tortora, M.; Cassani, G.; Fassina, G. *J. Mol. Recognit*. **1998**, *11*, 243-246.

(130) Palombo, G.; Rossi, M.; Cassani, G.; Fassina, G. *J. Mol. Recognit*. **1998**, *11*, 247-249.

(131) Palombo, G.; Verdoliva, A.; Fassina, G. *J. Chromatogr. B*. **1998**, *715*, 137-145.

(132) Fassina, G.; verdoliva, A.; Palombo, G.; Ruvo, M.; Cassani, G. *J. Mol. Recognit*. **1998**, *11*, 128-133.

(133) Fassina, G. In *Affinity Chromatography: Methods and Protocols*.; Bailon, P., Ehrlich, G. K., Fung, W. J., Berthold, W., Eds.; Humana Press: Totowa, New Jersey, 2000, p 57-68.

(134) Teng, S. F.; Sproule, K.; Husain, A.; Lowe, C. R. *J. Chromatogr. B*. **2000**, *740*, 1-15.

(135) *Handbook of Affinity Chromatography*; Kline, T. Ed.; Marcel Dekkler: New York, 1993.

(136) *Affinity Chromatography, Methods and Protocols*; Bailon, P., Ehrlich, G. K., Fung, W. J., Berthold, W., Eds.; Humana Press: Totowa, New Jersey, 2000.

(137) Fassina, G.; Ruvo, M.; Palombo, G.; Verdoliva, A.; Marino, M. *J. Biochem. Biophys. Methods* **2001**, *49*, 481-490.

(138) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531-1534.

(139) Keifer, P. A. *Drugs Future* **1998**, *23*, 301-317.

(140) Moore, J. M. *Curr. Opin. Biotechnol.* **1999**, *10*, 54-58.

(141) Diercks, T.; Coles, M.; Kessler, H. *Curr. Opin. Chem. Biol.* **2001**, *5*, 285-291.

(142) Stockman, B. J.; Farley, K. A.; Angwin, D. T. *Methods in Enzymology*. **2001**, *338*, 230-246.

(143) Meyer, B.; Peters, T. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 864-890.

(144) Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784-1788.

(145) Evans, J. N. S. *Biomolecular NMR Spectroscopy*; Oxford University Press: Oxford, 1995.

(146) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.

(147) Bystrov, V. F. *Prog. Nucl. Magn. Reson. Spectrosc.* **1976**, *10*, 41-81.

(148) Pardi, A.; Billeter, M.; Wuthrich, K. *J. Mol. Biol.* **1984**, *180*, 741-751.

(149) Karplus, M. *J. Phys. Chem.* **1959**, *30*, 11-15.

(150) Pardi, A.; Wagner, G.; Wuthrich, K. *Eur. J. Biochem.* **1983**, *137*, 445-454.

(151) Kumar, A.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *J. Am. Chem. Soc.* **1981**, *103*, 3654-3658.

(152) Macurra, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95-117.

(153) Wider, G.; Macura, S.; Kumar, A.; Ernst, R. R.; Wuthrich, K. *J. Magn. Reson.* **1984**, *56*, 207-234.

(154) Bothner-By, A. A.; Stephen, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811-813.

(155) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *63*, 207-213.

(156) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1982**, *48*, 402-417.

(157) Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1987**, *109*, 607-609.

(158) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Nat. Acad. Sci. U.S.A.* **1951**, *37*, 205-211.

(159) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wycoff, H.; Philips, D. C. *Nature* **1958**, *181*, 662-666.

(160) Blundell, T. L.; Barlow, D.; Barkakoti, N.; Thornton, J. M. *Nature* **1983**, *306*, 281-283.

(161) Chakarabarti, P.; Bernard, M.; Rees, D. C. *Biopolymers* **1986**, *25*, 1087-1093.

(162) MacArthur, W. M.; Thornton, J. M. *J. Mol. Biol.* **1996**, *264*, 1180-1195.

(163) Barlow, D.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601-619.

(164) Wuthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715-740.

(165) Clore, G. M.; Gronenborn, A. M. *Crit. Rev. Biochem. Mol. Biol.* **1989**, *24*, 479-564.

(166) Perzel, A.; Hollosi, M. In *Circular Dichroism and the Conformatinal Analysis of Biomolecules; Fasman*; Fasman, G. D., Ed.; Plenum Press: New York, 1996, p 362-364.

(167) Manning, M. C.; Illangasekare, M.; Woody, R. W. *Biophs. Chem.* **1988**, *31*, 77-86.

(168) Bush, C. A.; Sarkar, S. K.; Kopple, K. D. *Biochemistry* **1978**, *17*, 4951-4954.

(169) Woody, R. W.; Tinoco, I. J. *J. Chem. Phys.* **1967**, *46*, 4927-4945.

(170) Iizuka, E.; Yang, J. T. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1175.

(171) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108-4116.

(172) Woody, R. W. *Circular Dichroism Principles and Applications*; Nakanishi, K., Berova, N., Woody, R. W., Eds.; VCH: New York, 1994, p 473-496.

(173) Gans, P. J.; Lyu, P. C.; Manning, M. C.; Woody, R. W.; Kallenbach, N. R. *Biopolymers* **1991**, *31*, 1605-1614.

(174) Nelson, J. W.; Kallenbach, N. R. *Proteins* **1986**, *1*, 211-217.

(175) Sonnichsen, F. D.; Van Eyk, J. E.; Hodges, R. S.; Sykes, B. D. *Biochemistry* **1992**, *31*, 8790-8798.

(176) Baumann, H.; Hard, T. *Eur. J. Biochem.* **1995**, *230*, 879-885.

(177) Llinas, M.; Klein, M. P. *J. Am. Chem. Soc.* **1975**, *97*, 4731-4737.

(178) Lau, S. Y. M.; Taneja, A. K.; Hodges, R. S. *J. Chromatogr.* **1984**, *317*, 129-140.

(179) Albert, J. S.; Hamilton, A. D. *Biochemistry* **1995**, *34*, 984-990.

(180) Siedlecka, M.; Goch, G.; Ejchart, A.; Sticht, H.; Bierzyski, A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 903-908.

(181) Ghadiri, M. R.; Choi, C. *J. Am. Chem. Soc.* **1990**, *112*, 1630-1632.

(182) Ruan, F.; Chen, Y.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9403-9404.

(183) Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 9668-9676.

(184) Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470-2474.

(185) Yu, C.; Taylor, J. W. *Bioorg. Med. Chem.* **1999**, *7*, 161-175.

(186) Zhang, W.; Taylor, J. W. *Tetrahedron Lett.* **1996**, *37*, 2173-2176.

(187) Houston, M. E., Jr.; Gannon, C. L.; Kay, C. M.; Hodges, R. S. *J. Pept. Sci.* **1995**, *1*, 274-282.

(188) Bracken, C.; Gulyas, J.; Taylor, J. W.; Baum, J. *J. Am. Chem. Soc.* **1994**, *116*, 6431-6432.

(189) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1990**, *112*, 6046-6051.

(190) Phelan, J. C.; Skelton, N. J.; Braisted, A. C.; McDowell, R. S. *J. Am. Chem. Soc.* **1997**, *119*, 455-460.

(191) Jackson, D. Y.; King, D. S.; Chmielewski, J.; Singh, S.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 9391-9392.

(192) Ravi, A.; Prasad, B. V. V.; Balaram, P. *J. Am. Chem. Soc.* **1983**, *105*, 105-109.

(193) Toniolo, C.; Benedetti, E. *Macromolecules* **1991**, *24*, 4004-4009.

(194) Karle, I. L.; Flippen-Anderson, J. L.; Sukumar, M.; Uma, K.; Balaram, P. *J. Am. Chem. Soc.* **1991**, *113*, 3952-3956.

(195) Uma, K.; Balaram, P. *Ind. J. Chem.* **1989**, *28B*, 705-710.

(196) Marshall, G. R.; Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrocki, J.; Leplawy, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 487-491.

(197) Toniolo, C.; Bonora, G. M.; Bavoso, A.; Benedetti, E.; Blasio, B. D.; Pavone, V.; Pedone, C. *Biopolymers* **1983**, *22*, 205-215.

(198) Cammers-Goodwin, A.; Allen, T. J.; Oslick, S. L.; McClure, K. F.; Lee, J. H.; Kemp, D. S. *J. Am. Chem. Soc.* **1996**, *118*, 3082-3090.

(199) Serrano, L.; Fersht, A. R. *Nature* **1989**, *342*, 296-299.

(200) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Nature* **1987**, *326*, 563-567.

(201) Kallenbach, N. R.; Gong, Y. *Bioorg. Med. Chem.* **1999**, *7*, 143-151.

(202) Zhou, H. X.; Lyu, P. C.; Wemmer, D. E.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1994**, *116*, 1139-1140.

(203) Lyu, P. C.; Liff, M. I.; Marky, L. A.; Kallenbach, N. R. *Science* **1990**, *250*, 669-673.

(204) Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161-214.

(205) Patek, M. *Int. J. Pept. Protein Res.* **1993**, *42*, 97-117.

(206) Atherton, E.; Sheppard, R. C. *Solid Phase Peptide Synthesis, A Practical Approach*; IRL Press: Oxford, 1989.

(207) Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillessen, D. *Tetrahedron Lett.* **1989**, *30*, 1927-1930.

(208) Carpino, L. A.; Sadat-Aalaee, D.; Chao, H. G.; DeSelms, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 9651-9652.

(209) Carpino, L. A.; El-Faham, A. *J. Am. Chem. Soc.* **1995**, *117*, 5401-5402.

(210) Wenschuh, H.; Beyermann, M.; Winter, R.; Bienert, M.; Ionescu, D.; Carpino, L. A. *Tetrahedron Lett.* **1996**, *37*, 5483-5486.

(211) Rance, M.; Sfrensen, O. W.; BodenHausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479-85.

(212) Griesinger, C.; Otting, G.; Wuthrich, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1988**, *110*, 7870-7872.

(213) Pettitt, B. M.; Matsunaga, T.; Al-Obeidi, F.; Gehrig, C.; Hruby, V. J.; Karplus, M. *Biophy. J. Biophys. Soc.* **1991**, *60*, 1540-1544.

(214) Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 795-817.

(215) Achari, A.; Hale, S. P.; Howard, A. J.; Clore, G. M.; Gronenborn, A. M.; Hardman, K. D.; Whitlow, M. *Biochemistry* **1992**, *31*, 10449-10457.

(216) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliand, G. L. *Biochemistry* **1994**, *33*, 4721-4729.

(217) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906-918.

(218) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657-661.

(219) Lian, L. Y.; Derrick, J. P.; Sutcliffe, M. J.; Yang, J. C.; Roberts, G. C. K. *J. Mol. Biol.* **1992**, *228*, 1229-1234.

(220) Sauer-Eriksson, A. E.; Kleywegt, G. J.; Uhlen, M.; Jones, T. A. *Structure* **1995**, *3*, 265-278.

(221) Akerstrom, B.; Brodin, T.; Reis, K.; Bjorck, L. *J. Immunol.* **1985**, *135*, 2589-2592.

(222) Bjorck, L.; Kronvall, G. *J. Immunol.* **1984**, *133*, 969-974.

(223) Gronenborn, A. M.; Clore, G. M. *J. Mol. Biol.* **1993**, *233*, 331-335.

(224) Frick, I.-M.; Wikstrom, M.; Forsen, S.; Drakenberg, T.; Gomi, H.; Sjobring, U.; Bjorck, L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8532-8536.

(225) Akerstrom, B.; Bjorck, L. *J. Biol. Chem.* **1986**, *261*, 10240-10247.

(226) Reis, K. J.; Ayoub, E. M.; Boyle, M. D. P. *J. Immunol.* **1984**, *132*, 3098-3102.

(227) Reis, K. J.; Ayoub, E. M.; Boyle, M. D. *J. Immunol.* **1984**, *132*, 3091-3097.

(228) Stone, G. C.; Sjobring, U.; Bjorck, L.; Sjoquist, J.; Barber, C. V.; Nardella, F. A. *J. Immunol.* **1989**, *143*, 565-570.

(229) Fahnestock, S. R.; Alexander, P.; Nagle, J.; Filpula, D. *J. Bacteriol.* **1986**, *167*, 870-880.

(230) Guss, B.; Eliasson, M.; Olsson, A.; Uhlen, M.; Frej, A. K.; Jornvall, H.; Flock, J. I.; Lindberg, M. *EMBO J.* **1986**, *5*, 1567-1575.

(231) Va, P.; Vranken, D. V. *The UCI Undergraduate Research Journal* **2001**, *4*, 61-67.

(232) Newkome, G. R.; Moorefield, C. N.; Vogtle, F. *Dendritic Molecules: Concepts, Syntheses and Perpectives*; VCH: Weinheim, Germany, 1996.

(233) Tomalia, D. A.; Naylor, A.; Goddard III, W. A. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 138-175.

(234) Tomalia, D. A.; Hedstrand, D. M.; Wilson, L. R. *Encyclopedia of Polymer Science and Engineering, Index Volume*; Wiley: New York, 1990, p 46-92.

(235) Frechet, J. M. *Science* **1994**, *263*, 1710-1715.

(236) Ardoin, N.; Astruc, D. *Bull. Soc. Chim. Fr.* **1995**, *132*, 875-909.

(237) Frechet, J. M.; Hawker, C. J.; Gitsov, I.; Leon, J. W. *J. Macromol. Sci., Pure Appl. Chem.* **1996**, *A33*, 1399-1425.

(238) Bosman, A. W.; Janssen, H. M.; Meijer, E. W. *Chem. Rev.* **1999**, *99*, 1665-1688.

(239) Sadler, K.; Tam, J. P. *J Biotechnol* **2002**, *90*, 195-229.

(240) Fantl, W. J.; Johnson, D. E.; Williams, L. T. *Annu. Rev. Biochem.* **1993**, *62*, 453-481.

(241) Heldin, C.-H. *Cell* **1995**, *80*, 213-223.

(242) Li, Y.; Dias, J. R. *Chem. Rev.* **1997**, *97*, 283-304.

(243) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2754-2794.

(244) Kim, R. M.; Manna, M.; Hutchins, S. M.; Griffin, P. R.; Yates, N. A.; Bernick, A. M.; Chapman, K. T. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10012-10017.

(245) Qureshi, S. A.; Kim, R. M.; Konteatis, Z.; Biazzo, D. E.; Motamedi, H.; Rodrigues, R.; Boice, J. A.; Calaycay, J. R.; Bednarek, M. A.; Griffin, P.; Gao, Y. D.; Chapman, K.; Mark, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 12156-12161.

(246) Bo, Z.; Schafer, A.; Franke, P.; Schluter, D. *Org. Lett.* **2000**, *2*, 1645-1648.

(247) Zhang, W.; Nowlan III, D. T.; Thomson, L. M.; Lackowski, W. M.; Simanek, E. E. *J. Am. Chem. Soc.* **2001**, *123*, 8914-8922.

(248) Sivanandan, K.; Vutukuri, D.; Tayumanavan, S. *Org. Lett.* **2002**, *4*, 3751-3753.

(249) Zeng, F.; Zimmerman, S. C. *Chem. Rev.* **1997**, *97*, 1681-1712.

(250) Swali, V.; Wells, N. J.; Langley, J.; Bradley, M. *J. Org. Chem.* **1997**, *62*, 4902-4903.

(251) Cardona, C. M.; Jannach, S. H.; Huang, H.; Itojima, Y.; Leblanc, R. M.; Gawley, R. E. *Helv. Chim. Acta* **2002**, *85*, 3532-3558.

(252) Benters, R.; Neimeyer, C. M.; Wohrle, D. *ChemBiochem* **2001**, *2*, 686-694.

(253) Fromont, C.; Bradley, M. *Chem. Commun.* **2000**, 283-284.

(254) Li, Z.; Scheraga, H. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611-6615.

(255) Kirkpatrick, S.; Gelatt, C. D.; Jr & Vecchi, M. P. *Science* **1983**, *220*, 671-680.

(256) Luty, B. A.; Wasserman, Z. R.; Stouten, P. F. W.; Hodge, C. N.; Zacharias, M.; McCammon, J. A. *J. Comp. Chem.* **1995**, *16*, 454-464.

(257) Zelikin, A.; Shastri, V. R.; Langer, R. *J. Org. Chem.* **1999**, *64*, 3379-3380.

(258) Aquino, C. J.; Armour, D. R.; Berman, J. M.; Birkemo, L. S.; Carr, R. A.; Croom, D. K.; Dezube, M.; Dougherty, R. W., Jr.; Ervin, G. N.; Grizzle, M. K.; Head, J. E.; Hirst, G. C.; James, M. K.; Johnson, M. F.; Miller, L. J.; Queen, K. L.; Rimele, T. J.; Smith, D. N.; Sugg, E. E. *J. Med. Chem.* **1996**, *39*, 562-569.

(259) Majo, V. J.; Perumal, P. T. *J. Org. Chem.* **1998**, *63*, 7136-7142.

(260) Honigfort, M. E.; Brittain, W. J. *Macromolecules* **2002**, *35*, 4849-4851.

(261) Tsai, J. H.; Takaoka, L. R.; Powell, N. A.; Nowick, J. S. *Org. Syn.* **2002**, *78*, 220-224.

(262) *NovaBiochem Catalog & Peptide Synthesis Handbook*; Calbiochem Novabiochem Corp.: San Diego, 1999.

(263) Hauske, J. R.; Dorff, P. *Tetrahedron Lett.* **1995**, *36*, 1589-1592.

(264) Dressman, B. A.; Spangle, L. A.; Kaldor, S. W. *Tetrahedron Lett.* **1996**, *37*, 937-940.

(265) Roussel, P.; Bradley, M. *Tetrahedron Lett.* **1997**, *38*, 4861-4864.

(266) Horne, W. S.; Stout, C. D.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2003**, *125*, 9372-9376.

(267) Lober, S.; Rodriguez-Loaiza, P.; Gmeiner, P. *Org. Lett.* **2003**, *5*, 1753-1755.

(268) Lee, L. V.; Mitchell, M. L.; Huang, S.-J.; Fokin, V. V.; Sharpless, K. B.; Wong, C.-H. *J. Am. Chem. Soc.* **2003**, *125*, 9588-9589.

(269) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.

(270) Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064.

(271) Lai, K. O.; Fu, W. Y.; Ip, F. C.; Ip, N. Y. *Mol. Cell Neurosci.* **1998**, *11*, 64-76.

(272) Bibel, M.; Barde, Y. A. *Genes Dev*. **2000**, *14*, 2919-2937.

(273) Eide, F. F.; Lowenstein, D. H.; Reichardt, L. F. *Exp. Neurol*. **1993**, *121*, 200-214.

(274) McInnes, C.; Sykes, B. D. *Biopolymers* **1997**, *43*, 339-366.

(275) Gotz, R.; Koster, R.; Winkler, C.; Raulf, F.; Lottspeich, F.; Schartl, M.; Thoenen, H. *Nature* **1994**, *372*, 266-269.

(276) Patapoutian, A.; Reichardt, L. F. *Curr. Opin. Neurobiol*. **2001**, *11*, 272-280.

(277) Lee, F. S.; Kim, A. H.; Khursigara, G.; Chao, M. V. *Curr. Opin. Neurobiol*. **2001**, *11*, 281-286.

(278) Kaplan, D. R.; Hempstead, B. L.; Martin-Zanca, D.; Chao, M. V.; Parada, L. F. *Science* **1991**, *252*, 554-558.

(279) Lamballe, F.; Klein, R.; Barbacid, M. *Cell* **1991**, *66*, 967-979.

(280) Barbacid, M. *J. Neurobiol*. **1994**, *25*, 1386-1403.

(281) Chao, M. V.; Hempstead, B. L. *Trends Neurosci*. **1995**, *18*, 321-326.

(282) Ibanez, C. F. *Trends Neurosci*. **1998**, *21*, 438-444.

(283) Drinkwater, C. C.; Barker, P. A.; Suter, U.; Shooter, E. M. *J. Biol. Chem*. **1993**, *268*, 23202-23207.

(284) Ibánez, C. F.; Ebendal, T.; Barbany, G.; Murray-Rust, J.; Blundell, T. L.; Persson, H. *Cell* **1992**, *69*, 329-341.

(285) Ibánez, C. F.; Ebendal, T.; Persson, H. *EMBO J*. **1990**, *9*, 1477-1483.

(286) Ibánez, C. F.; Ilag, L. L.; Murray-Rust, J.; Persson, H. *EMBO J*. **1993**, *12*, 2281-2293.

(287) Kahle, P.; Burton, L. E.; Schmelzer, C. H.; Hertel, C. *J. Biol. Chem*. **1992**, *267*, 22707-22710.

(288) Rydén, M.; Murray-Rust, J.; Glass, D. I., L.L.; Trupp, M.; Yancoupoulos, G. D.; McDonald, N. Q.; Ibánez, C. F. *EMBO J*. **1995**, *14*, 1979-1990.

(289) Shih, A.; Laramee, G. R.; Schmelzer, C. H.; Burton, L. E.; Winslow, J. W. *J. Biol. Chem.* **1994**, *269*, 27679-27686.

(290) Urfer, R.; Tsoulfas, P.; Soppet, D.; Escandon, E.; Parada, L. F.; Presta, L. G. *EMBO J.* **1994**, *13*, 5896-5909.

(291) Rydén, M.; Ibánez, C. F. *J. Biol. Chem.* **1996**, *271*, 5623-5627.

(292) Kruttgen, A.; Jr., J. V. H.; Kahle, P. J.; Shooter, E. M. *J. Biol. Chem.* **1997**, *272*, 29222-29228.

(293) Kullander, K.; Kaplan, D.; Ebendal, T. *J. Biol. Chem.* **1997**, *272*, 9300-9307.

(294) Ibánez, C. F.; Ebendal, T.; Persson, H. *EMBO J.* **1991**, *10*, 2105-2110.

(295) Ilag, L. L.; Lonnerberg, P.; Persson, H.; Ibánez, C. F. *J. Biol. Chem.* **1994**, *269*, 19941-19946.

(296) Kullander, K.; Ebendal, T. *J. Neurosci. Res.* **1994**, *39*, 195-210.

(297) Urfer, R.; Tsoulfas, P.; O'Connel, L.; Presta, L. G. *Biochemistry* **1997**, *36*, 4775-4781.

(298) Wiesmann, C.; Ultsch, M. H.; Bass, S. H.; de Vos, A. M. *Nature* **1999**, *401*, 184-188.

(299) Salehi, A.; Berhaagen, J.; Dijkhuizen, P. A.; Swaab, D. F. *Neurosci.* **1996**, *75*, 373-387.

(300) Narisawa-Saito, M.; Wakabayashi, K.; Tsuji, S.; Takahashi, H.; Nawa, H. *Neuroreport.* **1996**, *7*, 2985-2988.

(301) Hefti, F. *J. Neurobiol.* **1994**, *25*, 1418-1435.

(302) Apfel, S. C.; Kessler, J. A. *Baillieres Clinical Neur.* **1995**, *4*, 593-606.

(303) Ebadi, M.; Bashir, R. M.; Heidrick, M. L.; Hamada, F. M.; Refaey, H. E.; Hamed, A.; Helal, G.; Baxi, M. D.; Cerutis, D. R.; Lassi, N. K. *Neurochem. Int'l.* **1997**, *30*, 347-374.

(304) Muragaki, Y.; Chou, T. T.; Kaplan, D. R.; Trojanowski, J. Q.; Lee, V. M.-Y. J. *Neurosci.* **1997**, *17*, 530-542.

(305) Washiyama, K.; Muragaki, Y.; Rorke, L. B.; Lee, V. M.; Feinstein, S. C.; Radeke, M. J.; Blumberg, D.; Kaplan, D. R.; Trojanowski, J. Q. *Am. J. Path.* **1996**, *148*, 929-940.

(306) Conner, J. M.; Tuszynski, M. H. *Mental Retardation and Developmental Diabilities Research Reviews* **1998**, *4*, 212-222.

(307) Hughes, R. A.; O'Leary, P. D. *Clin. Exp. Pharm. Physi.* **1996**, *23*, 956-959.

(308) Barinaga, M. *Science* **1994**, *264*, 772-774.

(309) Lindsay, R. M. *Ciba Foundation Symposium* **1996**, *196*, 39-53.

(310) Wang, W.; Yang, J.; Ying, J.; Xiong, C.; Zhang, J.; Cai, C.; Hruby, V. J. *J. Org. Chem.* **2002**, *67*, 6353-6360.

(311) Eguchi, M.; Shen, R. Y. W.; Shea, J. P.; Lee, M. S.; Kahn, M. *J. Med. Chem.* **2002**, *45*, 1395-1398.

(312) Burgess, K. *Acc. Chem. Res.* **2001**, *34*, 826-835.

(313) Eguchi, M.; Lee, M. S.; Nakanishi, H.; Stasiak, M.; Lovell, S.; Kahn, M. *J. Am. Chem. Soc.* **1999**, *121*, 12204-12205.

(314) Virgilio, A. A.; Bray, A. A.; Zhang, W.; Trinh, L.; Snyder, M.; Morrissey, M. M.; Ellman, J. A. *Tetrahedron* **1997**, *53*, 6635-6644.

(315) Pitt, N. G., D. *Tetrhedron Lett.* **1999**, *40*, 3811-3814.

(316) Kitagawa, O.; Velde, D. V.; Dutta, D.; Morton, M.; Takusagawa, F.; Aubé, J. *J. Am. Chem. Soc.* **1995**, *117*, 5169-5178.

(317) Bach, A. C.; Espina, J. R.; Jackson, S. A.; Stouten, P. F. W.; Duke, J. L.; Mousa, S. A.; Degrado, W. F. *J. Am. Chem. Soc.* **1996**, *118*, 293-294.

(318) Roedern, E. G. V.; Lohof, E.; Hessler, G.; Hoffmann, M.; Kessler, H. *J. Am. Chem. Soc.* **1996**, *118*, 10156-10167.

(319) Feng, Y.; Wang, Z.; Jin, S.; Burgess, K. *J. Am. Chem. Soc.* **1998**, *120*, 10768-10769.

(320) Park, G.; Burgess, K. *J. Comb. Chem.* **2001**, *3*, 257-266.

(321) Feng, Y.; Burgess, K. *Biotech. Bioeng. Comb. Chem.* **1999**, *71*, 3-8.

(322) Reyes, S.; Pattarawarapan, M.; Roy, S.; Burgess, K. *Tetrahedron* **2000**, *56*, 9809-9818.

(323) Li, W.; Burgess, K. *Tetrahedron Lett.* **1999**, *40*, 6527-6530.

(324) Feng, Y.; Wang, Z.; Jin, S.; Burgess, K. *Chem.-Eur. J.* **1999**, *5*, 3273-3278.

(325) Burgess, K.; Feng, Y.; Pattarawarapan, M.; Wang, Z. *J. Org. Chem.* **1999**, *64*, 9175-9177.

(326) Feng, Y.; Pattarawarapan, M.; Wang, Z.; Burgess, K. *Org. Lett.* **1999**, *1*, 121-124.

(327) Maliartchouk, S.; Feng, Y.; Ivanisevic, L.; Debeir, T.; Cuello, A. C.; Burgess, K.; Saragovi, H. U. *Mol. Pharm.* **2000**, *57*, 385-391.

(328) Pattarawarapan, M.; Reyes, S.; Xia, Z.; Zaccaro, M. C.; Saragovi, H. U.; Burgess, K. *J. Med. Chem.* **2003**, *46*, 3565-3567.

(329) Berteina, S.; Mesmaeker, A. D. *Synlett.* **1998**, 1227-1230.

(330) Mutter, M.; Hersperger, R. *Synthesis* **1989**, 198-200.

(331) Hudlicky, M. *Reductions in Organic Chemistry*; Wiley: New York, 1984.

(332) Rink, H. *Tetrahedron Lett.* **1987**, *28*, 3787-3790.

(333) Neustadt, B. R.; Smith, E. M.; Nechuta, T.; Zhang, Y. *Tetrahedron Lett.* **1998**, *39*, 5317-5320.

(334) Lee, H. B.; Zaccaro, M. C.; Pattarawarapan, M.; Roy, S.; Saragovi, H. U.; Burgess, K. *J. Org. Chem.* **2004**, *69*, 701-13.

(335) Yong, Y. F.; Kowalski, J. A.; Lipton, M. A. *J. Org. Chem.* **1997**, *62*, 1540-1542.

(336) Maliartchouk, S.; Saragovi, H. U. *J. Neurosci.* **1997**, *17*, 6031-6037.

APPENDIX A

A SCRIPT FOR MINIMIZATION, DYNAMICS AND CUTOFF IN DISCOVER FOR CHAPTER II

```
#BIOSYM btcl 3
#
#  Input File For Discover Generated By Zebin Xia
#  Date:          Tue Jul  3 20:52:22 2001
#  User Name: xia
#  Host Name: Burgess1
#  Host Type: iris
#
#  System Name: KH2_S1
#
#Stage Name: 1  begin
  set PROJECT kkk
  set FORCEFIELD /chem/insight/I2000/irix6m4/biosym_lib/cff/cff.frc
  begin
  readFile coordinate filename = $PROJECT.car
#
#Stage Name: 2  nonbonds
  forcefield nonbond \
    +separate_coulomb \
    vdw \
      summation_method = no_cutoff \
    coulomb \
      -distance_dependent_dielectric \
      dielectric_value = 45 \
     summation_method = no_cutoff
  #
#
#Stage Name: 3  minimize
  minimize \
    iteration_limit = 1000 movement_limit = 0.200 \
    sd \
      convergence = 10.0 line_search_precision = 0.100 \
    cg \
      convergence = 1.0 method = polak \
      line_search_precision = 0.100 \
     newton \
```

```
     convergence = 0.001 method = newton_raphson \
     line_search_precision = 0.900 max_atoms = 200 \
   final_convergence = 0.001
  #
set step mini
  writeFile coordinate filename = $PROJECT$step.cor
#
set temp_start 10.0
set temp_final 1000.0
set iterations 1000
set num 100

#Stage Name: 4 dynamics
   #
if {$num <= 0} return
    if {$num == 1} {
        $temp = $temp_final
    } else {
        $temp = $temp_start
    }
    for {$i = 0} {$i < $num} {incr i} {
        $temp = $temp_start + $i * 10.0
        dynamics \
        time = 160. timestep = 1.0 \
        initial_temperature = $temp-5.0 -boltzman \
        ensemble = nvt temperature_control_method = velocity_scaling \
        integration_method = Velocity_verlet \
        temperature = $temp temperature_window = 10 \
        deviation = 2000.000
    }
set step heat
  writeFile coordinate filename = $PROJECT$step.cor
#
#Stage Name: 5  dynamics
  rattle bonds -tolerance 1e-5
#
  dynamics \
    time = 12000 timestep = 1.0 \
    initial_temperature = 1000.0 -boltzmann \
    ensemble = nve \
    deviation = 5000
  #
set step equil
  writeFile coordinate filename = $PROJECT$step.cor
```

```
#
#Stage Name: 6  dynamics
  dynamics \
    time = 600000 timestep = 1.0 \
    initial_temperature = 1000.0 -boltzmann \
    ensemble = nvt temperature_control_method = velocity_scaling \
    integration_method = Velocity_verlet \
    temperature = 1000.0 temperature_window = 10 \
    deviation = 50000 \
    execute frequency = 1000 last_step = 0 +after \
      command = {print archive +coordinates \
                filename = $PROJECT.arc}
set step prod
  writeFile coordinate filename = $PROJECT$step.cor

 #
#end input file for minimization and dynamics
#



#
  begin
#
#Stage Name: 2  nonbonds
  forcefield nonbond \
    +separate_coulomb \
    vdw \
      summation_method = no_cutoff \
    coulomb \
      -distance_dependent_dielectric \
      dielectric_value = 45 \
      summation_method = no_cutoff
#
#Stage Name: 3  loop for minimization
   for {$frame = 1} {$frame < 600} {$frame = $frame + 1} {
```

```
#
#Stage Name: 4  file control
   readFile archive filename = $PROJECT.arc frame = $frame
#
#Stage Name: 5  minimize
   minimize \
    iteration_limit = 3000 movement_limit = 0.200 \
    sd \
      convergence = 10.0 line_search_precision = 0.100 \
    cg \
      convergence = 1.00 method = polak \
      line_search_precision = 0.100 \
    newton \
      convergence = 0.001 method = bfgs \
      line_search_precision = 0.900 max_atoms = 200 \
    final_convergence = 0.001

 set step mini
#
#Stage Name: 6  file control
   writeFile archive filename = $PROJECT$step.arc frame = $frame
#
#Stage Name: 7  end loop
  }

#end input file for the minimization of 600 structures



#
```

```
# value to add to the lowest energy
set value 3.6

 begin

# set parameters for energy calculations

 forcefield nonbond \
  +separate_coulomb \
  vdw \
    summation_method = no_cutoff \
  coulomb \
    -distance_dependent_dielectric \
    dielectric_value = 45 \
    summation_method = no_cutoff

# find lowest energy structure
#

set step 0
set cut_off 0.0
set frame_lowest 0

for {$i = 1} {[set energy [readFile archive filename = $PROJECT$step.arc frame = $i]]
!= ""} {incr i} {
  if {$energy <= $cut_off} {
     $cut_off = $energy
     $frame_lowest = $i
  }
```

```
}
#
# print the lowest energy value to the output file
echo "This is the lowest energy and its frame number in 600"
echo $cut_off
echo $frame_lowest

echo "This is the cutoff value"
$cutoff_value = $cut_off+$value

# print the cut off value to the output file
echo $cutoff_value

set step1 cutoff
echo ""
echo "Starting the archive search"

$j = 0

#
# pull out the structures that meet the cut off criteria and
# save them to $PROJECTcutoff.arc

for {$i = 1} {[set energy [readFile archive filename = $PROJECT$step.arc frame = $i]]
!= ""} {incr i} {
# print the energy of each structure to the output file
# there should be 600
   echo $energy
   if { $energy <= $cutoff_value } {
```

```
      $j = $j+1
      echo "last energy meets cut off criteria"
# calculate the energy and save the coordinates and E to archive file
      energy
      writeFile archive filename = $PROJECT$step1.arc frame = $j
  if { $energy == $cut_off } {
     echo "This is the lowest energy and its frame number in the cutoffed list"
     echo $energy
     echo $j
   }
 }
}

#end input file for cutoff
```

APPENDIX B

PROCEDURES FOR QMD STUDIES WITH INSIGHT II/DISCOVER FOR CHAPTER II

Starting InsightII

Open **UNIX shell**

Type **i2k**

Step 1: Building a **3D** Molecule

Go to Module

select **Builder**

or select **Molbuilder** from **Toobox**

A. Draw the **2D** molecule

In somecases, such as **NH3+,** after drawing the molecule, go to Atom/Charge, have **Formal_Charge** be on, type **1**, then execute.

Go to sketcher, and put the file .car2d.

Click 2D 3D

If stereochemistry is not correct, go to Toolbox/stereochemistry.

Go to **Moleculer**, and **rename** the file .car.

Or directly click Fragment Librtaries to build a 3D molecule.

Step 2: Select Forcefield and Potentials

Go to **FF/potentials**

Potential Action

- Fix

Partial Chg Action

- Fix

Formal Chg Action

- Accept

Click execute

Go to **FF/Assign_CFF**

Potential Action

- Fix

Partial Chg Action

- Fix

Formal Chg Action

- Accept

Click execute

Step3: Refine the model
Click **Optimize**

Open **Discovery_3**

Step 4: Run our own program (**dyna.inp**)
This program is for the ***Minimization***, ***Heating*** and ***Equilibration.***

Step 5: Run our own program (**mini.inp**)
This file is for the ***minimization of 600 structures***.

Step 6: Run our own program (**cutoff.inp**)
This file is for the **cut off,** like 114 structures from 600 structures.

Step 7: Build a cluster graph
Open **Analysis**
Go to **Trajectory/get** (pick up your cutoff.arc file)
Trajectory/conformation (Pick the lowest energy structure)
Then go to **Trajectory/Cluster_Graph**
Choose: ***Add*** in List Control,
Specified in Atom Set,
Atom in molecule pick level.
Pick atom subset (eg. 14 atoms).
Have **End_Definition** be **ON**.
Set: **Max RMS Value**, eg. 1.32
Number RMS Levels, eg. 4

Step8: Build a family.
Go to **Trajectory/Family**
Fill out **Family Name** like F1
Define **Family Mode**
_Box
Structure Pair
Fill out with some frames, like 83,1
Structure Pair 2
Fill out with e.g. 83,114
Fill out **Value** with e.g. 0.0-0.6
Fill out **Tolerance** with e.g. 0.01
Choose **Reference structure**
_lowest_Energy
Click execute
Go to **Trajectory/put**

Select Mode
_Specified
Frame Spec
Fill out with e.g. 3,6-9…
_Sort_Frame
Archive File Name
Name it like F1.arc
Click execute

Repeat the above, we can get Family 2, 3, 4, and so on. .

Step 9: Measure Distance and Dihedral angles
Go to **Trajectory/Conformation**
Fill in the frame number with the lowest energy for each family, then execute.
Go to **Measure/Distance** or
Measure/Dihedral.

APPENDIX C

A SCRIPT FOR MINIMIZATION, DYNAMICS AND CUTOFF IN CHARMM FOR CHAPTER II

```
* Script file produced by Zebin Xia
*
! Startup script for CHARMm
!
UPPER  ! case for file to write
BOMBLEVEL -2
WRNLEV 0
PRNLEV 5
! Script to read parameter, psf and crd file
!
OPEN READ UNIT 21 CARD NAME
"/chem/insight/I2000/irix6m3/biosym_lib//AMINOH.RTF"
READ RTF UNIT 21 CARD
CLOSE UNIT 21
OPEN READ UNIT 21 CARD NAME
"/chem/insight/I2000/irix6m3/biosym_lib//CHARMM_support.rtf"
READ RTF UNIT 21 CARD APPEND
CLOSE UNIT 21

OPEN READ UNIT 20 CARD NAME
"/chem/insight/I2000/irix6m3/biosym_lib//PARM.PRM"
READ PARA UNIT 20 CARD
CLOSE UNIT 20

OPEN READ UNIT 20 CARD NAME "b8_1_2.psf"
READ PSF UNIT 20 CARD
CLOSE UNIT 20

OPEN READ UNIT 20 CARD NAME "b8_1_2.crd"
READ COOR UNIT 20 CARD
CLOSE UNIT 20
```

```
SKIPE EXCL BOND ANGL DIHE IMPR VDW ELEC HBON USER HARM CDIH
CIC CDRO NOE


MINIMIZE -
  SD NSTEP 1000 -
  STEP 0.020000 NPRI 5 -
  TOLENR 0.000000 TOLGRD 0.010000 TOLSTP 0.000000 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD

SKIPE EXCL BOND ANGL DIHE IMPR VDW ELEC HBON USER HARM CDIH
CIC CDRO NOE

MINIMIZE -
  ABNR NSTEP 3000 -
  STEP 0.020000 NPRI 5 -
  TOLENR 0.000000 TOLGRD 0.010000 TOLSTP 0.000000 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD


SHAKE BONH TOL 1e-09 MAXIT 500


UPDATE -
  INBFRQ -1 CUTNB 15.000000 CTONNB 11.000000 CTOFNB 14.000000 -
  VSWITCH SWITCH CDIE EPS 45.000000 -
  IHBFRQ 0 -
  IMGFRQ 20 CUTIM 15.000000 IXTFRQ 100 NOEWALD
  SHAKE BONH TOLR 1e-09 MXIT 500 PARAM
  SKIPE EXCL BOND ANGL DIHE IMPR VDW ELEC HBON USER HARM CDIH
CIC CDRO NOE



OPEN WRIT UNIT 31 CARD NAME "b8_heat_1.rst"
OPEN WRIT UNIT 32 FILE NAME "b8_heat_1.dcd"
OPEN WRIT UNIT 33 FILE NAME "b8_heat_1.vel"
```

```
OPEN WRIT UNIT 34 CARD NAME "b8_heat_1.ene"

DYNAmics VERL STRT -
  TIME 0.001000 NSTEP 12000 -
  FIRSTT 0.000000 FINALT 1000.000000 TEMINC 5.0 -
  ISEED 314159 -
  IHTFRQ 60 IEQFRQ 0 IASORS 1 IASVEL 1 ICHECW 0 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ -1 -
  ISVFRQ 20 NSAVC 20 NPRINT 20 -
  IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD -
  IUNREA -1 IUNWRI 31 IUNCRD 32 IUNVEL -1 KUNIT 34

OPEN READ UNIT 30 CARD NAME "b8_heat_1.rst"
OPEN WRIT UNIT 31 CARD NAME "b8_equil_1.rst"
OPEN WRIT UNIT 32 FILE NAME "b8_equil_1.dcd"
OPEN WRIT UNIT 33 FILE NAME "b8_equil_1.vel"
OPEN WRIT UNIT 34 CARD NAME "b8_equil_1.ene"

DYNAmics VERL REST -
  TIME 0.001000 NSTEP 12000 ISEED 314159 -
  FINALT 1000.000000 -
  TWINDH 10.000000 TWINDL -10.000000 -
  IHTFRQ 0 IEQFRQ 200 NTRFRQ 50 IASORS 0 ISCVEL 0 ICHECW 1 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD -
  ISVFRQ 20 NSAVC 20 NPRINT 20 -
  IUNREA 30 IUNWRI 31 IUNCRD 32 IUNVEL -1 KUNIT 34

OPEN READ UNIT 30 CARD NAME "b8_equil_1.rst"
OPEN WRIT UNIT 31 CARD NAME "b8_1.rst"
OPEN WRIT UNIT 32 FILE NAME "b8_1.dcd"
OPEN WRIT UNIT 33 FILE NAME "b8_1.vel"
OPEN WRIT UNIT 34 CARD NAME "b8_1.ene"

DYNAmics VERL REST -
  TIME 0.001000 NSTEP 600000 FINALT 1000.000000 -
  IHTFRQ 0 IEQFRQ 0 NTRFRQ 0 ICHECW 0 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
```

```
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD -
  ISVFRQ 1000 NSAVC 1000 NPRINT 1000 -
  IUNREA 30 IUNWRI 31 IUNCRD 32 IUNVEL -1 KUNIT 34


WRITE TITLE UNIT 6
* FINAL ENERGY = ?TOTK ?ENER

OPEN WRIT UNIT 41 CARD NAME "b8_1_final.crd"
WRIT COOR UNIT 41 CARD
* Coordinates after stage
*

OPEN READ UNIT 32 FILE NAME "b8_1.dcd"
OPEN WRIT UNIT 42 FILE NAME "b8_mini.dcd"
OPEN WRIT UNIT 34 CARD NAME "b8_energy.ene"
TRAJECTORY IREAD 32 SKIP 1 IWRITE 42
* minimized trajectory
*

SET 1 1
LABEL LOOP
TRAJ READ
MINIMIZE -
  SD NSTEP 1000 -
  STEP 0.020000 NPRI 5 -
  TOLENR 0.000000 TOLGRD 0.010000 TOLSTP 0.000000 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD

MINIMIZE -
  ABNR NSTEP 3000 -
  STEP 0.020000 NPRI 5 -
  TOLENR 0.000000 TOLGRD 0.010000 TOLSTP 0.000000 -
  INBFRQ -1 CUTNB 15.000000 -
  CTONNB 11.000000 CTOFNB 14.000000 VSWITCH SWITCH -
  CDIE EPS 45.000000 -
  IHBFRQ 0 IMGFRQ 20 CUTIM 15.000000 IXTTFRQ 100 NOEWALD
TRAJ WRITE
GETE
FORMAT (F16.4)
```

```
set 2 0.0000
set 4 ?ENER
INCREMENT 2 by 0.0
INCREMENT 4 by 0.0
FORMAT (I16)
set 3 @1
INCREMENT 3 by 0
WRITE TITLE UNIT 34
*@3@2@4@2@4
*@2@2@2@2@2
*@2@2@2@2@2
*@2@2

INCR 1 by 1
IF 1 LE ?NFILE GOTO LOOP


OPEN READ UNIT 41 FILE NAME "b8_mini.dcd"
OPEN WRIT UNIT 42 FILE NAME "b8_low.dcd"
OPEN WRIT UNIT 43 CARD NAME "b8_l0w.ene"
TRAJECTORY IREAD 41 SKIP 1 IWRITE 42
* trajectory created by selecting frames with lowest energy
*
set  2  1
set  8  100.0
set  1  1
set  7  1
set  9  2.6
label  loop
traj   read
energy CDIE EPS 45.0
if  8  gt  ?ener  set  8  ?ener  set  7  @2
if  9  lt  ?ener  goto  next
traj  write
write  title  unit 43
*@2            ?ener
incr  2  by  1
label  next
incr  1  by  1
if  1  LE  600  goto  loop
print  @7
print  @8

stop
```

APPENDIX D

PROCEDURES FOR QMD STUDIES WITH INSIGHT II/CHARMM FOR CHAPTER II

Starting InsightII

Open **UNIX shell**

Type **i2k**

Step 1: Building a **3D** Molecule

Go to Module

select **Builder**

or select **Molbuilder** from **Toobox**

A. Draw the **2D** molecule

In somecases, such as **NH3+,** after drawing the molecule, go to Atom/Charge, have **Formal_Charge** be on, type **1**, then execute.

Go to sketcher, and put the file .car2d.

B. Click 2D 3D

If stereochemistry is not correct, go to Toolbox/stereochemistry.

Go to **Moleculer**, and **rename** the file .crd and .psf.

C. Or directly click Fragment Librtaries to build a 3D molecule.

Step 2: Select Forcefield and Potentials

Go to **FF/Selecttion**

Choose CHARMm.rfrc

Go to **FF/potentials**

Potential Action

- Fix (sometimes, use accept)

Partial Chg Action

- Fix (sometimes, use accept)

Formal Chg Action

- Accept

Click execute

Go to **FF/Assign_CFF (If select cff as force field)**

Potential Action

- Fix

Partial Chg Action

- Fix

Formal Chg Action

- Accept

Click execute

Step3: Open **CHARMm**
Choose run/using existing_file (like b1.inp)

Step4: Go to Quanta for Analysis

APPENDIX E
EXPERIMENTAL FOR CHAPTER III

General Methods. All α-amino acids used were of the L-configuration. All chemicals were obtained from commercial suppliers and directly used without further purification. HypoGel 400 RAM Resin was obtained from Rapp Polymere and amino acids were obtained from Advanced Chem Tech. 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), *N*- hydroxybenzotriazole (HOBt), di-iso-propylethylamine(DIEA), trifluoroacetic acid (TFA), CH_2Cl_2, DMF, piperidine and tri-iso-propylsilane (TIS) were purchased from Aldrich. *N*,*N'*-Diisopropylcarbodiimide (DIC) was purchased from Chem-Impex international. Tetramethyl fluoro formamidinium hexafluorophosphate (TFFH) was prepared based on the known method (Boas, U; Pedersen, B; Christensen, J. B. *synthetic communications* **1998**, 28, 1223-1231. Dourtoglou, V.; Gross, B. *synthesis* **1984**, 572-574). DMF was stored over 4Å molecular sieves for a few days before use. After each solid phase reaction, the resin was washed with DMF(3x), MeOH(3x), CH_2Cl_2(3x). Reverse phase high performance liquid chromatography (RP-HPLC) was carried out using BACKMAN system on Vydac C-18 columns of the following dimensions: 25 x 2.2 cm for preparation, and 25 x 0.46 cm for analysis. MALDI-TOF mass spectra were recorded on a Voyager-Elite XL instrument (PerSeptive Biosystems, Framington, MA). All HPLC experiments were performed using gradient conditions. The eluants used were solvent A (H_2O with 0.1% TFA) and solvent B (CH_3CN with 0.1% TFA). Flow rates used were 10 mL/min for preparative HPLC, and 1.0 mL/min for analytical HPLC. All NMR spectra were recorded on varian instruments at 500 MHz. NMR chemical shifts are expressed in δ ppm relative to internal solvent peaks (H_2O: 4.78 ppm, DMSO-d6: 2.50 ppm), and coupling constants were measured in Hz. Multiplicities in ^1H NMR were reported as s (singlet), d (doublet), t (triplet), m (multiplet). The phosphate buffer (20 mM, PH=4) was prepared with

sodium dihydrogenphosphate monohydrate and pure water. 0.1 M HCl was used to adjust PH value. CD spectra were recorded on an Aviv model 62 DS spectrometer.

General Procedure for Syntheses of Peptide Sequences 1-7. Rink resin in a fritted syringe was swelled in CH_2Cl_2 for 30 min. The Fmoc protecting group on the resin was removed by treating it twice with 20 % piperidine in DMF (10min and 15 min). The resin was then washed using a typical washing cycle: DMF (3x), CH_3OH (3x) and CH_2Cl_2 (3x). The Fmoc-amino acid (4 eq.), HBTU(4 eq.), HOBt(4 eq.), and DIEA (6 eq.) in DMF were added. After 2 h of gentle shaking, a ninhydrin test on a small sample of beads gave a negative result. The resin was filtered and washed using the washing cycle described above. The Fmoc protecting group was then removed by treating the resin with 20 % piperidine in DMF (2x: 10 min and 15 min). This coupling /deprotection cycle was repeated to build the desired sequence, except the coupling when Fmoc-Aib-OH was used the coupling agent used was TFFH instead of HBTU and HOBt. Finally, the peptide was cleaved from the resin by a treatment with a mixture of 90% TFA, 5% TIS, and 5% H_2O for 2 h. The cleavage solution was separated from the resin by filtration, then most of the cleavage cocktail was evaporated in a stream of nitrogen, and the crude peptide was precipitated using cooled anhydrous ethyl ether. The purity of this crude product was determined by analytical HPLC (SSI system, 5-95% B in 30 min). The crude product was dissolved in H_2O, filtered, then purified via preparative HPLC (Beckman system) and finally lyophilized to yield the desired product.

CD Studies. The peptidomimetics were dissolved in buffered water and at various concentrations of added hexafluoroisopropanol as specified in the text. CD measurements were performed in a 0.1 cm pathlength cell at pH 4.0, 25 ˚C, 50 μM, averaging 5 scans with a step size of 0.5 nm for the spectra.

For monitoring if the addition of IgG to a synthesized peptide solution increases the helix content of the peptide, a special cell, which has two separated components, was used. The pathlength of each component is 0.437 cm. The stoichiometry of ligand with

IgG used in this study was 2:1 which means 2 mg ligand and 1 mg IgG in 1 ml pH 4.0 buffer. The following experiments were performed for each sequence at 25 ˚C and averaging 5 scans with a step size of 0.5 nm: (a) a pH 4.0 buffer solution in one component, IgG solution in the other component; (b) mixed the pH 4.0 buffer and IgG solution; (c) a peptide solution in one component, IgG solution in the other component; (d) mixed the peptide solution and IgG solution.

The following formula (Gans, P. J.; Lyu, P. C.; Woody, R. W.; Kallenbach, N. R. *Biopolymers* **1991**, 31, 1605-1614) was employed to estimate the helix content:

$$\% \text{ helix} = [\ -[\square]_{222}/(40{,}000(n-4)/n)] \times 100\%$$

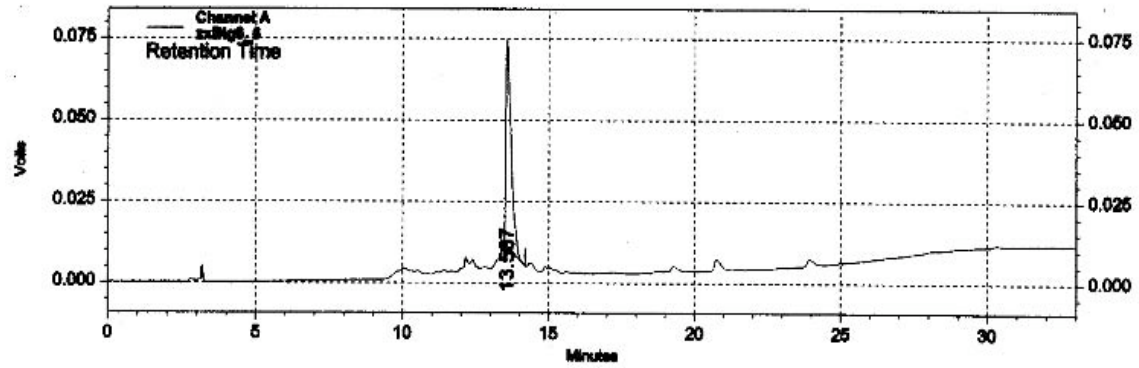
General Procedures for NMR Studies. All samples were prepared to give a 50 μM concentration of the peptidomimetic. Three kinds of solvent systems were used: DMSO-d6, 9:1 H_2O:D_2O and H_2O : $(CF_3)_2CDOD$ in the rations specified in the chapter VI. For the aqueous solvents, presaturation was carried out to suppress the H_2O signal. One-dimensional ^1H NMR spectra were recorded with a spectral width of 8000 Hz, 30272 data points, 16 transients, and a 3 s acquisition time. Some vicinal coupling constants were obtained from one-dimensional spectra at 25 ºC. Assignments of ^1H NMR resonances were performed using sequential connectivities.

ROESY or NOESY with mixing times of 150, 250, 300, 350, and 400 ms, DQF-COSY and TOCSY with mixing time of 80 ms spectra were recorded in the pure absorption mode according to the method of States et al (States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Res.* **1982**, 48, 286-292). All spectra were recorded with a 1.5 ~ 3 s relaxation delay, 512 t1 increments, and 16 or 32 scans per t1 increments with 2K data points at t2. Each 2D-data set was zero-filled to 2K x 2K data sets, and Gaussian transformed in both dimensions. All spectra were processed on a Silicon Graphics O_2 workstation using Varian (VNMR) and NMR Draw software. Sequential assignments of ^1H resonances , and structure determinations of peptidomimetics in solution were carried out by the following standard methods (Wuthrich, K. In *NMR of proteins and nucleic acids*, Wiley; New York, **1986**. Bax, A. *Annu. Rev. Biochem.* **1989**, 58, 223-256). The

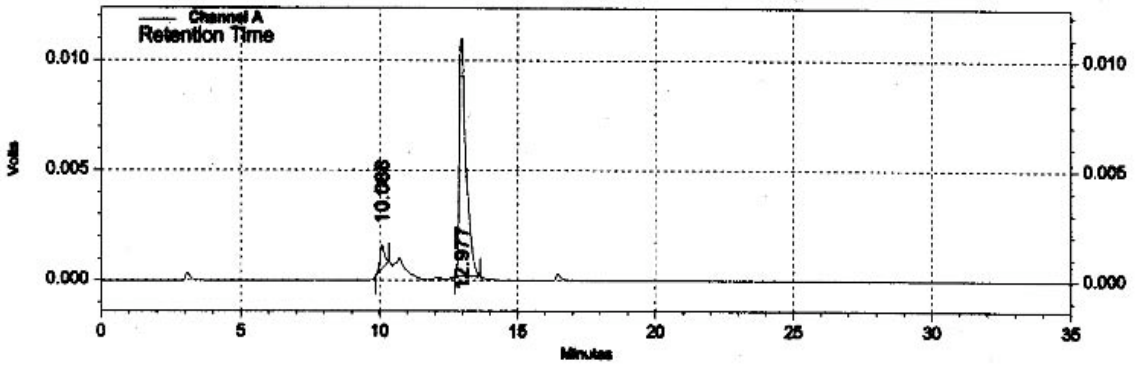
intensities of the ROESY or NOESY cross-peaks were assigned as S (strong), M(medium), and W (weak) based on the number of contours, and quantified as distance constraints of 2-3 Å (S), 3-4 Å (M), and 4-5 Å (W).(Gong, Y.; Zhou, H. X.; Guo, M.; Kallenbach, N. R. *Protein Sci*. **1995**, 4, 1446-1456) $^3J_{HN,H\alpha}$ coupling constants were measured from 1D NMR spectra.

Molecular Modeling. Calculations were performed on a Silicon Graphics O_2 workstation, operating under the IRIX 6.3 operating system using the modeling program CHARMm (version 23.2 Revision: 96.0501) with Quanta environment. Mechanics simulations were carried out using CHARMm forcefield. Biharmonic potential was used for dihedral angle constraints (E_ϕ) as well as for ROE or NOE distance constraints. Dihedral angle constraints to $-85 < \phi < -35$ for residues with $^3J_{HN\alpha} < 6$ Hz, and to $-175^\circ < \phi < -80^\circ$ for residues with $^3J_{HN\alpha} > 8$ Hz were used. The residue topology files (RTF) for all peptidomimetics were built using Quanta 2000 (vesion 2000, Molecular Simulations Inc.). Quenched molecular dynamics simulations for sequences were performed using the CHARMm standard parameters. All molecules were modeled in a dielectric constant of 80 (representing H_2O). The starting structure was minimized using 2000 steps of Steepest Descents (SD) and 2500 steps of the Adopted-Basis Newton Raphson method (ABNR) respectively until an RMS energy derivative of ≤ 0.01 kcal mol^{-1} A^{-1} was obtained. The minimized structure was then subjected to heating, equilibration, and dynamics simulation. The equations of motion were integrated using the Verlet algorithm with a time step of 1 fs, and SHAKE was used to constrain all bond lengths containing polar hydrogens. The minimized structure was heated to 1000K over 12 ps and equilibrated for another 12 ps at 1000K, then molecular dynamics runs were performed for a total time of 200 ps with trajectories saved every 1ps. The resulting 200 structures were thoroughly minimized again, using 2000 steps of SD followed by 3000 steps of ABNR. After excluding structures more than 6-7 kcal/mol above the lowest energy conformer identified, for peptidomimetics **1**, **4**, **5** and **6**, the remaining conformers were clustered into groups based on RMS deviation of 0.4, 0.8, 0.8, 0.3 for

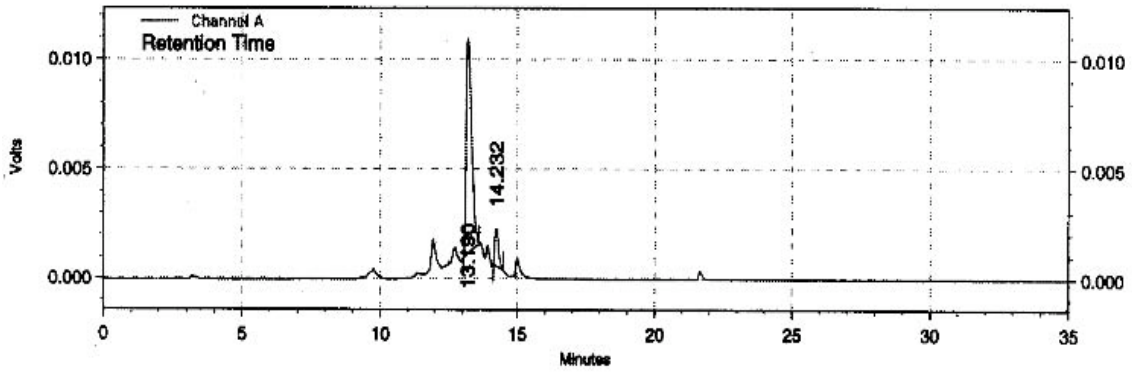
peptidomimetics **1**, **4**, **5** and **6** respectively, for a subset of main-chain backbone atoms excluding those associated with the *N*- and *C*-terminal amino acids.



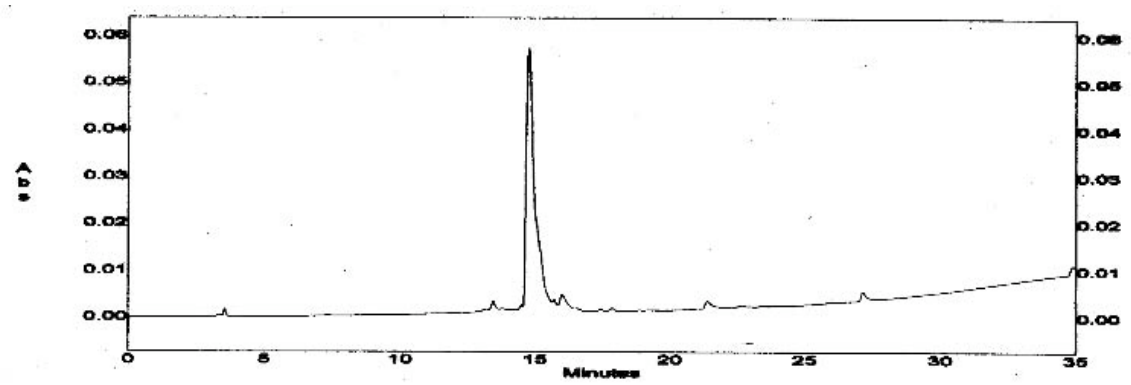
Crude HPLC trace of compound **1**



Crude HPLC trace of compound **4**



Crude HPLC trace of compound **5**



Crude HPLC trace of compound **6**

Table E1. Chemical Shifts and Coupling Constant Information for Peptidomimetic 1 (in 4:1 Buffer:$(CF_3)_2CDOD$)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.29/5.8	4.52	2.72, 2.72			
Phe-2	7.61/5.4	4.3	3.04, 3.04			3,4,5H:7.3-7.21(m), 2,6H:7.11 (d,J=6.0)
Tyr-3	7.46/4.0	4.94	3.02, 2.95			2,6H: 7.1(d,J=7.0), 3,5H: 6.64(d,J=7.0)
Aib-4	7.86		1.33, 1.23			
Ile-5	7.49/5.78	3.76	1.88	1.44, 1.20, 0.89	0.81	
Leu-6	7.73/4.95	4.08	1.56, 1.56	1.25	0.78, 0.73	
His-7	7.82/7.0	4.38	3.23, 2.75			2H: 8.48, 4H: 7.20
Acp-8	7.57/5.5					$(CH_2)_5$: 3.23, 3.14, 2.24, 1.52, 1.28
Gln-9	7.74/5.5	4.2	1.98, 1.88	2.29, 2.29		δNH_2: 7.22, 6.46
His-10	8.30/	4.63	3.22, 3.11			2H: 8.01, 4H: 6.96
Asn-11	7.87/5.47	4.59	2.85, 2.72			δNH_2: 7.18, 6.65
Aib-12	7.59		1.52, 1.42			
Phe-13	7.50/	4.44	3.17, 3.06			2,3,4,5,6H: 7.3-7.15(m)
Ile-14	7.62/	4.05	1.85	1.40, 1.13, 0.87	0.8	
Asn-15	7.99/5.9	4.69	2.82, 2.70			δNH_2: 7.32, 6.64
Succinyl						CH_2: 2.58, CH_2: 2.48, 2.41
NH_2	7.19, 6.85					

Table E2. Chemical Shifts and Coupling Constant Information for Peptidomimetic 2 (in DMSO – d6)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.40/6.5	4.48	2.64, 2.44			
Phe-2	8.0/7.5	4.3	3.0, 2.86			3,4,5H: 7.18-7.25(m), 2,6H: 7.13(d,J=7.5)
Tyr-3	7.90/6.0	4.14	2.98, 2.83			2,6H: 7.01(d,J=8.5), 3,5H: 6.64(d,J=8.5)
Aib-4	8.04		1.36, 1.28			
Ala-5	7.59/5.5	3.95	1.27			
Leu-6	7.50/7.0	4.02	1.53, 1.53	1.28	0.78, 0.72	
His-7	7.67/8.5	4.43	3.15, 2.70			2H: 8.94, 4H: 7.11
Acp-8	7.43/5.5					$(CH_2)_5$: 3.05, 2.06, 1.44, 1.35, 1.19
Ala-9	8.01/6.5	4.17	1.16			
His-10	8.21/8.0	4.58	3.06, 2.93			2H: 8.94, 4H: 7.32
Asn-11	8.20/5.5	4.45	2.54, 2.54			δNH_2: 7.72, 7.26
Aib-12	8.53		1.25, 1.04			
Phe-13	8.03/7.0	4.25	3.16, 2.93			3,4,5H: 7.19-7.27(m), 2,6H: 7.22(d,J=7.5)
Ile-14	7.54/8.5	4.09	1.78	1.47, 1.16, 0.83	0.73	
Asn-15	7.92/8.0	4.46	2.57, 2.41			δNH_2: 7.40, 6.94
Succinyl						CH_2: 2.43, CH_2: 2.33
NH_2	7.08, 6.86					

Table E3. Chemical Shifts and Coupling Constant Information for Peptidomimetic 3 (in 9:1 Buffer:D_2O)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.58/6.5	4.61	2.79, 2.70			
Phe-2	8.22/5.5	4.47	3.05, 2.99			3,4,5H:7.51-7.42(m), 2,6H:7.28(m)
Tyr-3	8.09/4.6	4.25	3.02, 2.98			2,6H: 7.28(m), 3,5H: 6.95(d,J=7.0)
Aib-4	8.27		1.41, 1.38			
Ala-5	7.91/4.8	4.18	1.38			
Leu-6	7.90/	4.14	1.56, 1.56	1.37	0.87, 0.81	
His-7	8.23/7.0	4.56	3.25, 2.90			2H: 8.72, 4H: 7.06
Acp-8	7.9					$(CH_2)_5$: 3.18, 3.18, 2.24, 1.47, 1.26
Ala-9	8.05/	4.24	1.37			
Arg-10	8.53/4.8	4.27	1.78, 1.78	1.61, 1.61	3.18, 3.18	NH: 7.39 (t, J=5.4)
Asn-11	8.56/5.5	4.6	2.84, 2.75			δNH_2: 7.76, 6.91
Aib-12	8.43		1.44, 1.43			
Ala-13	8.38/4.3	4.25	1.35			
Ile-14	8.04/	4.12	1.9	1.37, 1.21, 0.92	0.86	
Asn-15	8.42/6.5	4.68	2.85, 2.71			δNH_2: 7.81, 6.94
Succinyl						CH_2: 2.64, CH_2: 2.51
NH_2	7.58, 7.14					

Table E4. Chemical Shifts and Coupling Constant Information for Peptidomimetic 4 (in 17:3 Buffer:$(CF_3)_2CDOD$)

Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.28/5.9	4.66	2.90, 2.90			
Gln-2	8.49/4.5	4.1	2.03, 2.03	2.32, 2.32		δNH_2: 7.29, 6.68
Gln-3	8.14/4.5	4.15	2.09, 2.09	2.36, 2.36		δNH_2: 7.29, 6.65
Asn-4	7.95/	4.57	2.87, 2.77			δNH_2: 7.16, 6.64
Ala-5	7.98/3.85	4.11	1.4			
Phe-6	7.92/5.1	4.33	3.15, 3.15			3,4,5H:7.25-7.2(m), 2,6H:7.13 (d,J=7.5)
Tyr-7	8.0/3.9	3.93	3.19, 3.14			2,6H: 7.07(d,J=8.0), 3,5H: 6.61(d,J=8.0)
Aib-8	8.41		1.57, 1.45			
Ala-9	8.05/3.80	3.97	1.47			
Leu-10	7.78/4.85	4.04	1.46, 1.46	1.38	1.14, 0.64	
His-11	7.8/7.0	4.34	3.26, 2.71			2H: 7.83 (s), 4H: 6.97 (s)
Acp-12	7.79/					$(CH_2)_5$: 3.28, 3.13, 2.27, 1.54, 1.31
Ala-13	7.72/4.78	4.2	1.36			
Arg-14	8.02/4.8	4.15	1.79, 1.79	1.62, 1.62	3.15, 3.15	NH: 7.07
Asn-15	7.94/	4.57	2.83, 2.77			δNH_2: 7.18, 6.65
Ala-16	7.87/5.3	4.09	1.38			
Aib-17	7.94		1.44, 1.39			
Ile-18	7.36/7.0	3.99	1.91	1.49, 1.21, 0.88	0.84	
Gln-19	7.95/	4.21	2.13, 2.03	2.39, 2.39		δNH_2: 7.36, 6.70
Asn-20	8.07/7.0	4.69	2.85, 2.72			δNH_2: 7.03, 6.64
Succinyl						CH_2: 2.67, CH_2: 2.56
NH_2	7.29, 6.97					

Table E5. Chemical Shifts and Coupling Constant Information for Peptidomimetic 5 (in 17:3 Buffer:$(CF_3)_2CDOD$)

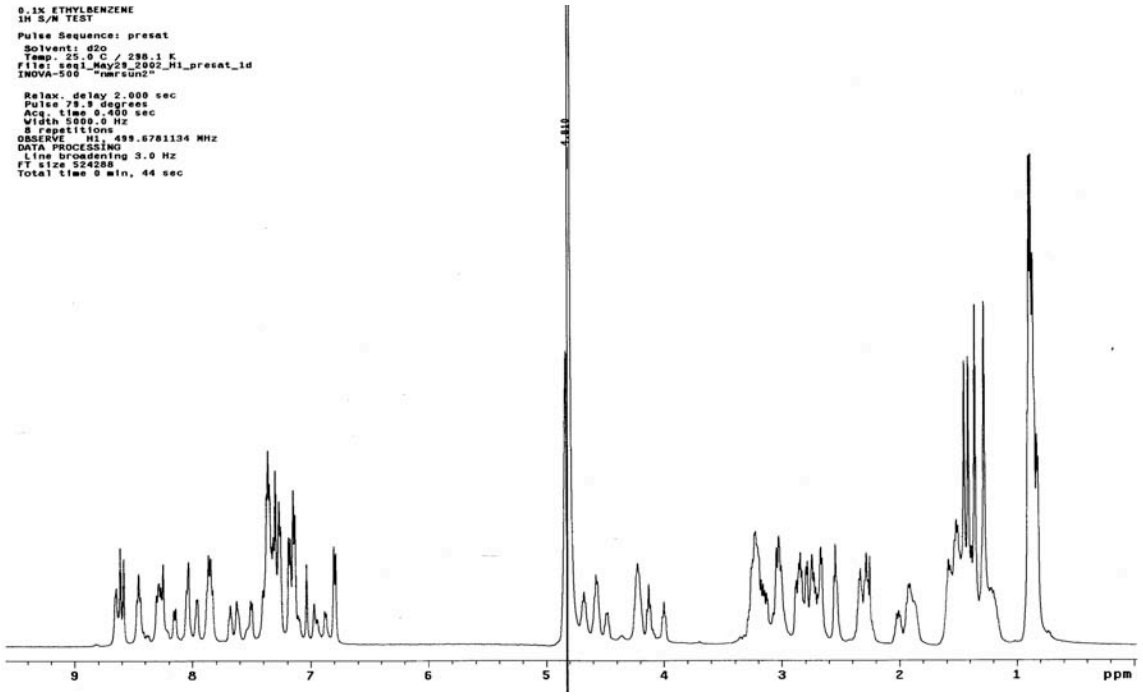
Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp-1	8.19/5.2	4.52	2.87, 2.82			
Ala-2	8.04/3.85	4.1	1.35			
Phe-3	7.47/5.0	4.33	3.08, 3.08			3,4,5H:7.32-7.25(m), 2,6H:7.09 (d,J=8.5)
Tyr-4	7.76/4.8	3.92	3.12, 3.00			2,6H: 7.04(d,J=8.5), 3,5H: 6.61(d,J=8.5)
Aib-5	8.39		1.52, 1.44			
Ala-6	8.86/3.85	3.94	1.45			
Leu-7	7.71/	4.03	1.80, 1.80	1.44	1.12, 0.63	
His-8	7.79/	4.54	3.25, 2.71			2H: 7.86, 4H: 6.96
Acp-9	7.79/					$(CH_2)_5$: 3.27, 3.10, 2.27, 1.53, 1.29
Ala-10	7.73/4.0	4.03	1.41			
Arg-11	7.98/5.8	4.03	1.76, 1.76	1.63, 1.63	3.12, 3.12	NH: 6.99
Asn-12	7.83/	4.44	2.87, 2.78			δNH_2: 7.54, 6.79
Ala-13	7.69/4.15	4.16	1.35			
Aib-14	8.02		1.46, 1.40			
Ile-15	7.67/6.5	3.74	1.88	1.39, 1.26, 0.88	0.82	
Gln-16	7.89/3.92	3.96	2.16, 2.16	2.43, 2.43		δNH_2: 7.20, 6.61
Ser-17	7.94/	4.28	4.03, 3.92			
Leu-18	7.92/	4.16	1.83, 1.83	1.51	0.82, 0.82	
Lys-19	8.09/6.5	4.11	1.84, 1.84	1.45, 1.45	1.64, 1.64	δCH_2: 2.94, 2.94, NH_2: 7.55
Asn-20	7.82/	4.67	2.87, 2.78			δNH_2: 7.06, 6.79
Succinyl						CH_2 2.69, 2.65, CH_2 2.56, 2.54
NH_2	7.23, 7.04					

Table E6. Chemical Shifts and Coupling Constant Information for Peptidomimetic 6 (in 17:3 Buffer:$(CF_3)_2CDOD$)

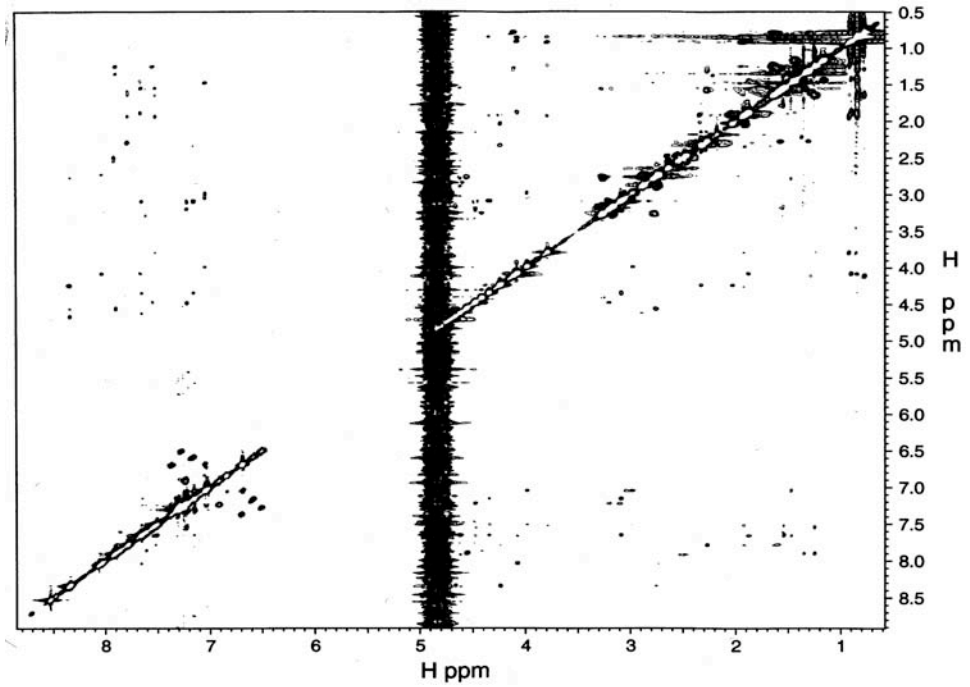
Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Asp	7.92/5.5	4.58	2.83, 2.83			
Phe	7.62/4.15	4.29	3.03, 3.03			3,4,5H:7.26-7.19(m), 2,6H:7.09(d,J=7.0)
Tyr	7.52/3.85	3.91	2.95, 2.95			2,6H: 6.98(d,J=7.5), 3,5H: 6.64(d,J=7.5)
Aib	7.6		1.47, 1.41			
Ile	7.42/4.9	3.73	1.86	1.58, 1.16, 0.85	0.77	
Leu	7.78/4.8	4.03	1.54, 1.54	1.25	0.77, 0.73	
His	7.92/5.9	4.35	3.21, 2.68			2H: 8.17, 4H: 6.97
Succinyl						CH_2: 2.57, CH_2: 2.47, 2.41
NH_2	7.11, 6.88					

Table E7. Chemical Shifts and Coupling Constant Information for Peptidomimetic 7 (in DMSO – d6)

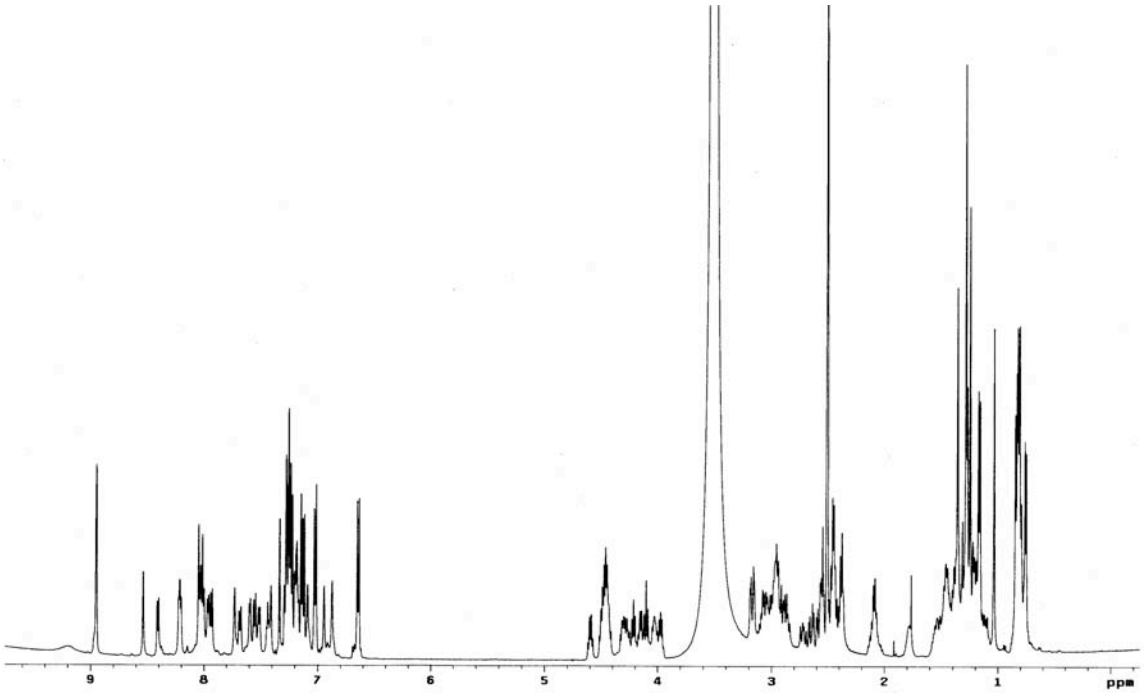
Residue	$\delta(H_N)/^3J_{HN\alpha}$	$\delta(H_\alpha)$	$\delta(H_\beta)$	$\delta(H_\gamma)$	$\delta(H_\delta)$	Others
Gln-1	8.10/7.0	4.15	1.83, 1.68	2.10, 2.10		δNH_2: 7.26, 6.78
His-2	8.22/7.5	4.62	3.05, 2.94			2H: 8.94, 4H: 7.30
Asn-3	8.25/7.0	4.49	2.57, 2.57			δNH_2: 7.72, 7.24
Aib-4	8.56		1.25, 1.03			
Phe-5	8.04/7.5	4.27	3.17, 2.94			2,3,5,6H: 7.29-7.2(m), 4H: 7.2-7.14(m)
Ile-6	7.53/8.0	4.11	1.79	1.43, 1.12, 0.84	0.77	
Asn-7	7.91/8.0	4.47	2.57, 2.41			δNH_2: 7.39, 6.93
NH_2	7.07, 6.85					
CH_3CO-						CH_3: 1.84



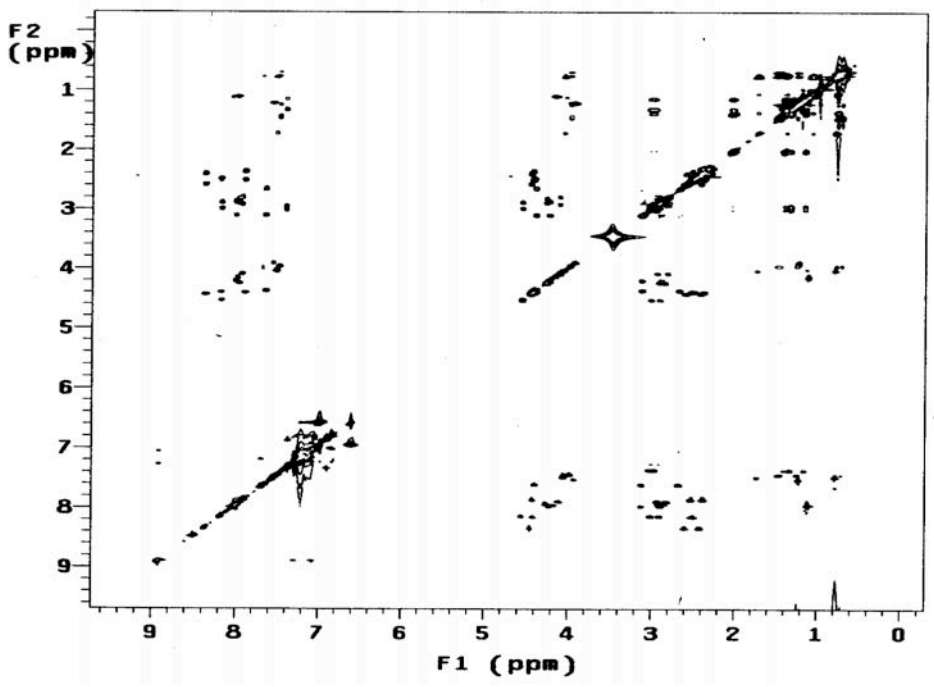
^{1}H NMR of peptidomimetic **1** (90% H_2O/10% D_2O)



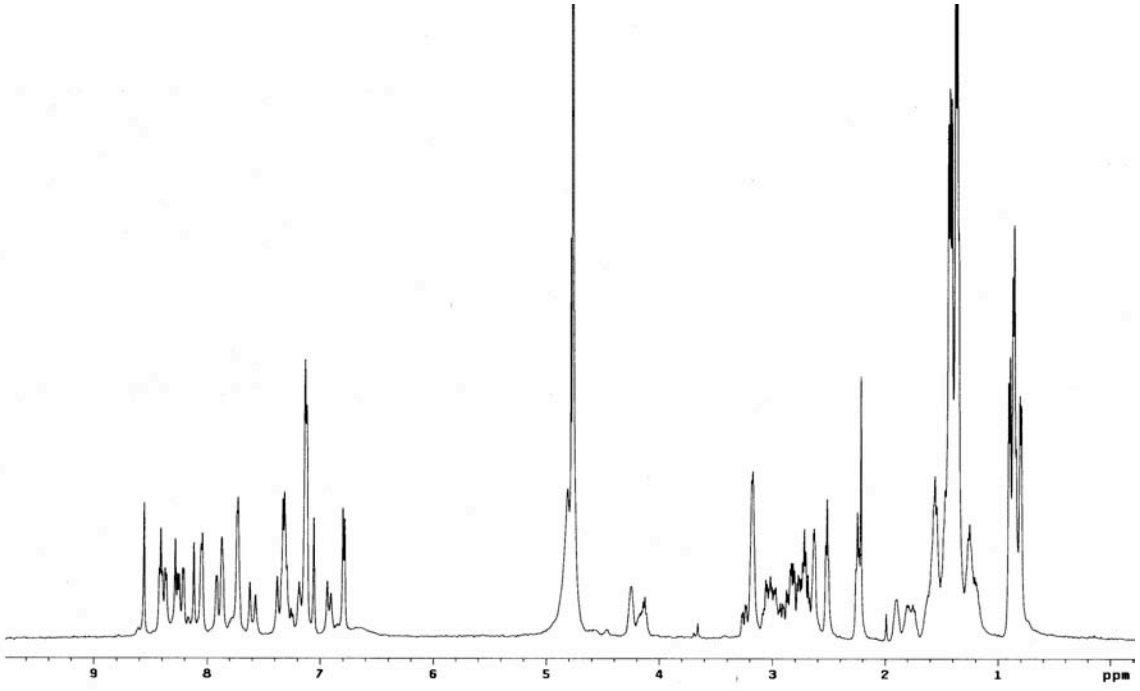
200-ms ROESY of peptidomimetic **1** (85% H_2O/20% $(CF_3)_2CDOD$)



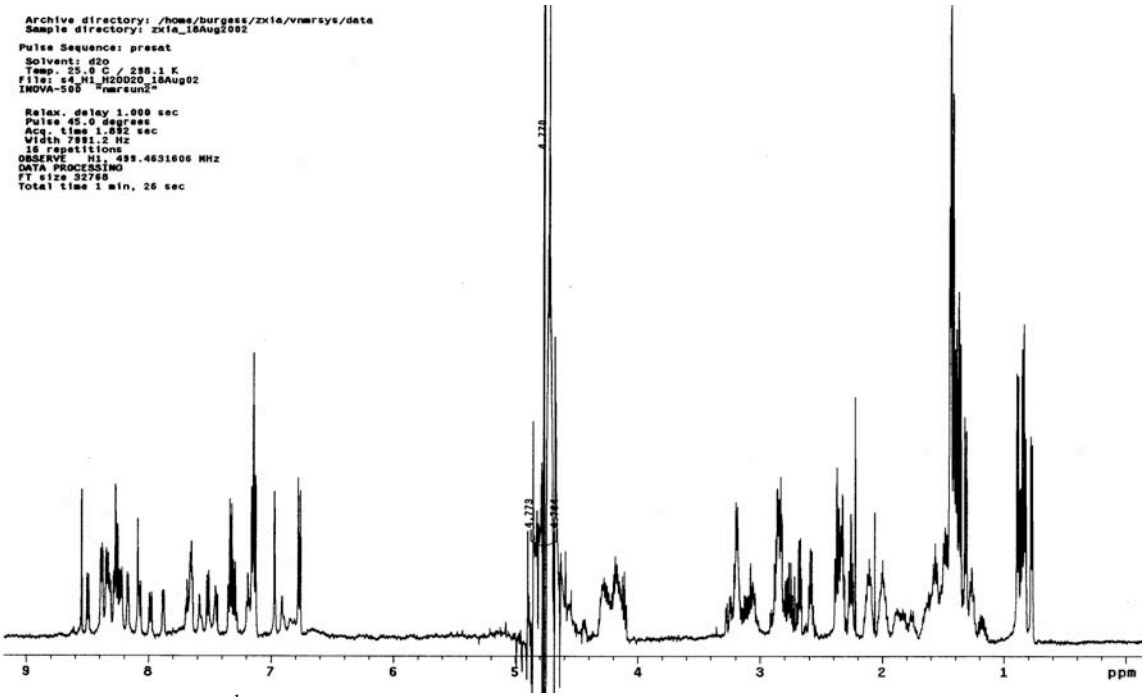
^{1}H NMR of peptidomimetic **2** (DMSO-d6)



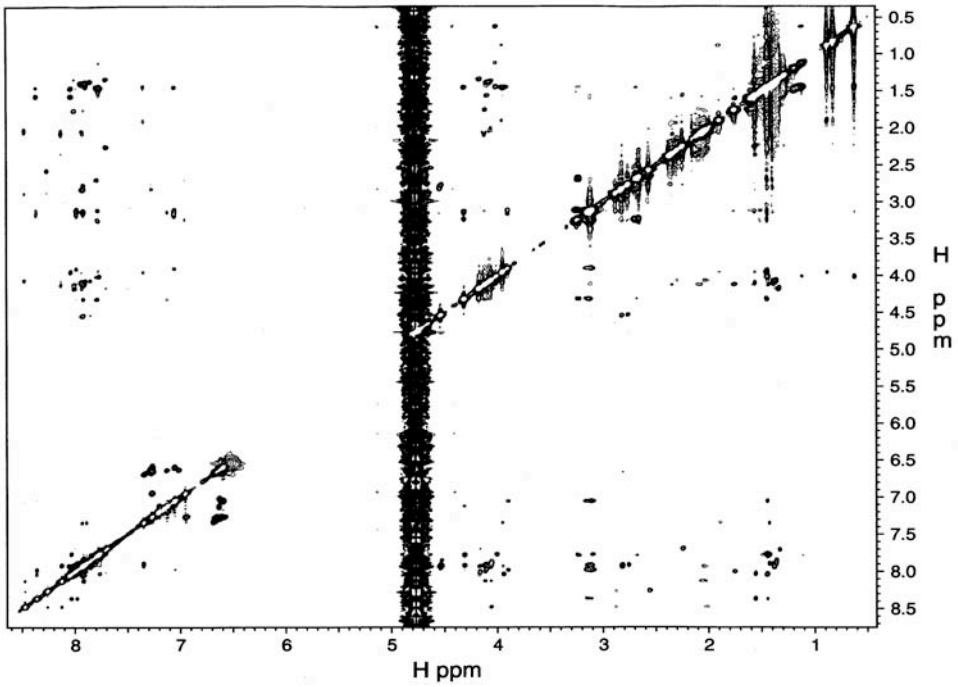
80-ms TOCSY of peptidomimetic **2** (DMSO-d6)



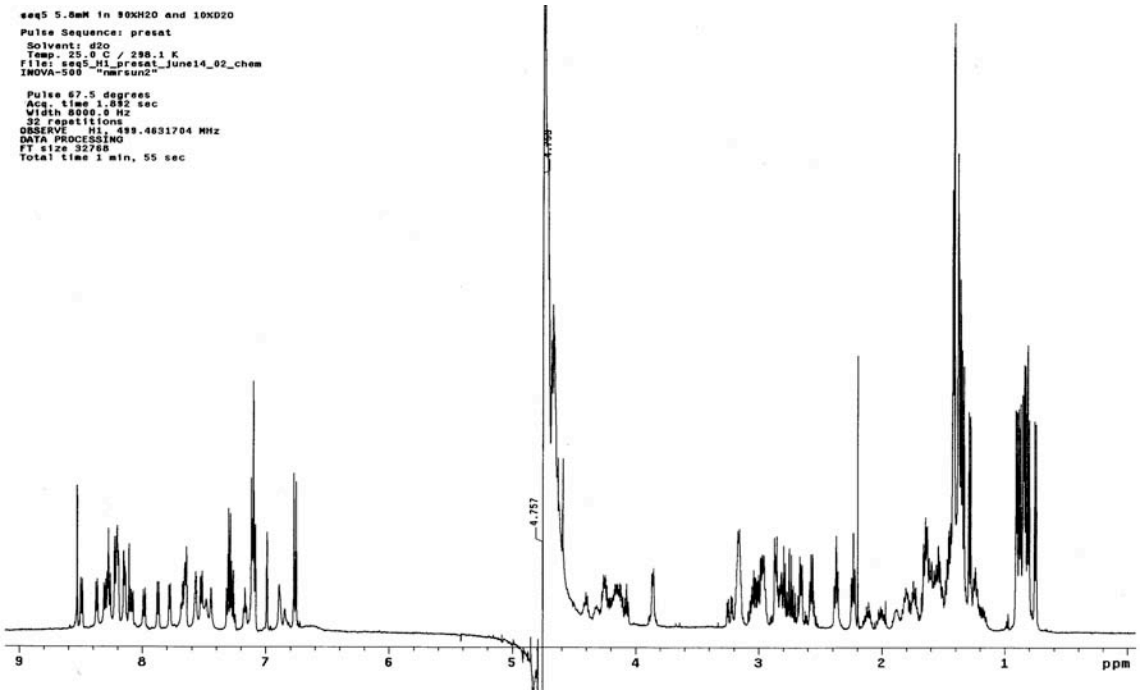
^{1}H NMR of peptidomimetic **3** (90% H_2O/10% D_2O)



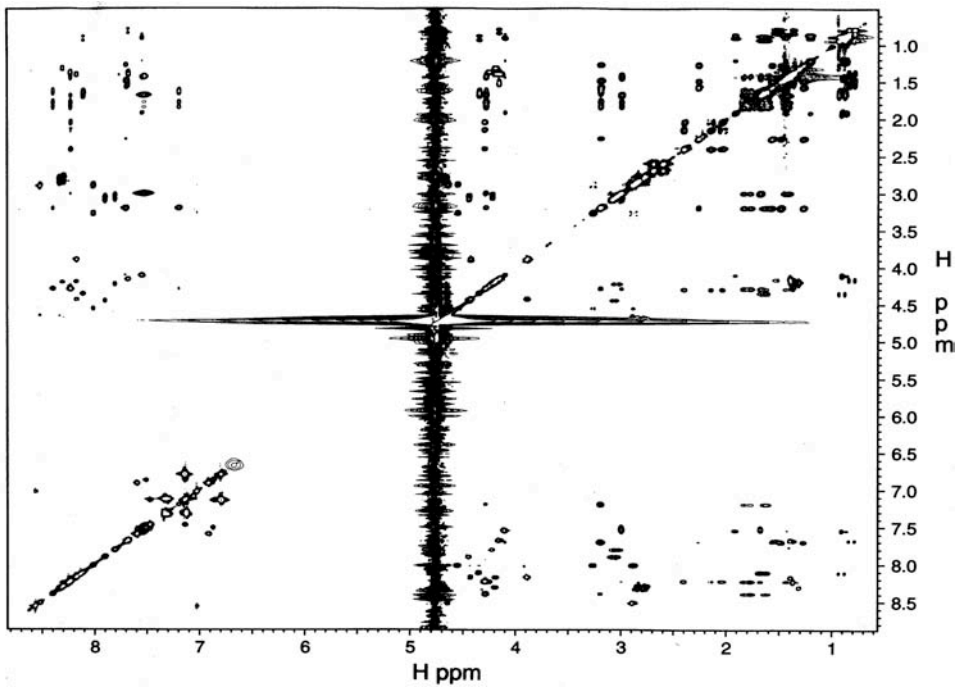
^{1}H NMR of peptidomimetic **4** (90% H_2O/10% D_2O)



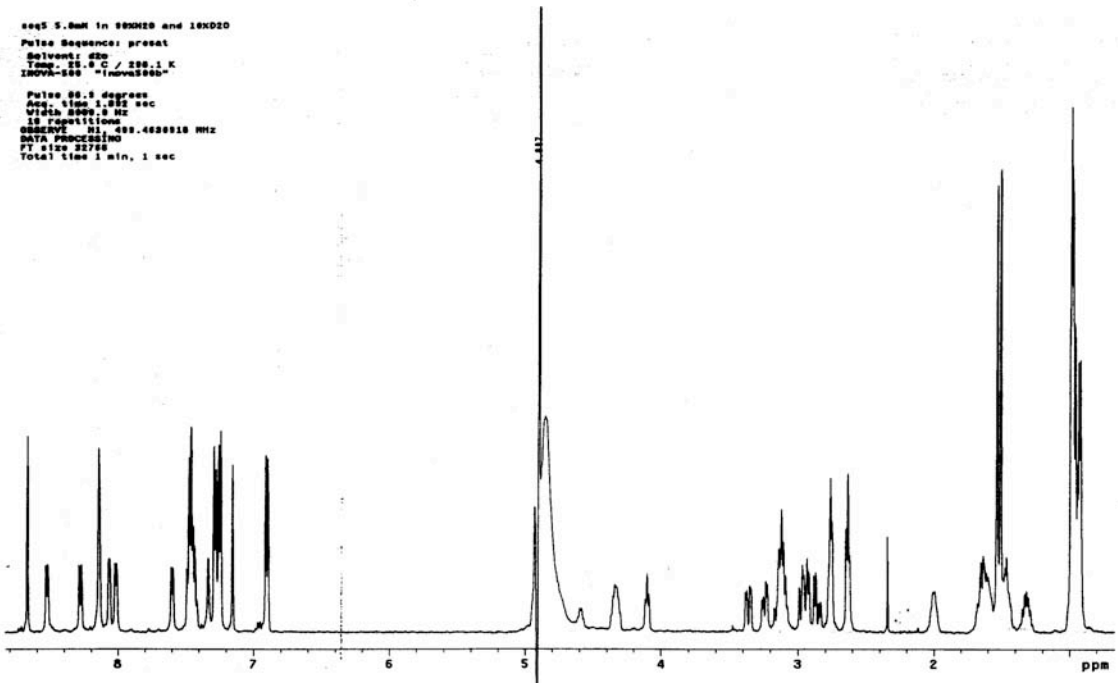
250-ms NOESY of peptidomimetic **4** (85% H_2O/15% $(CF_3)_2CDOD$)



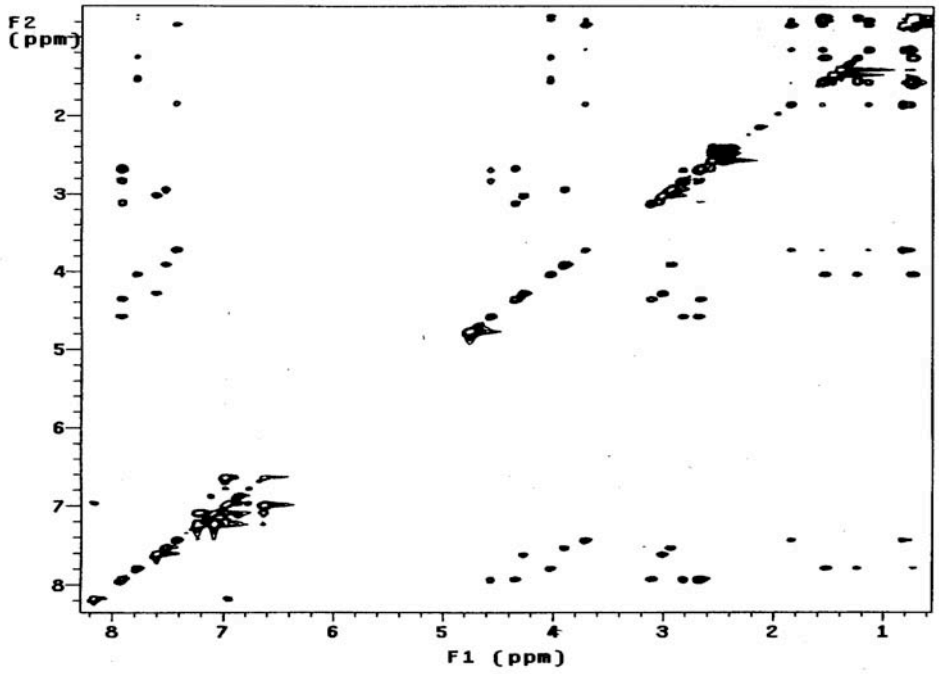
^{1}H NMR of peptidomimetic **5** (90% H_2O/10% D_2O)



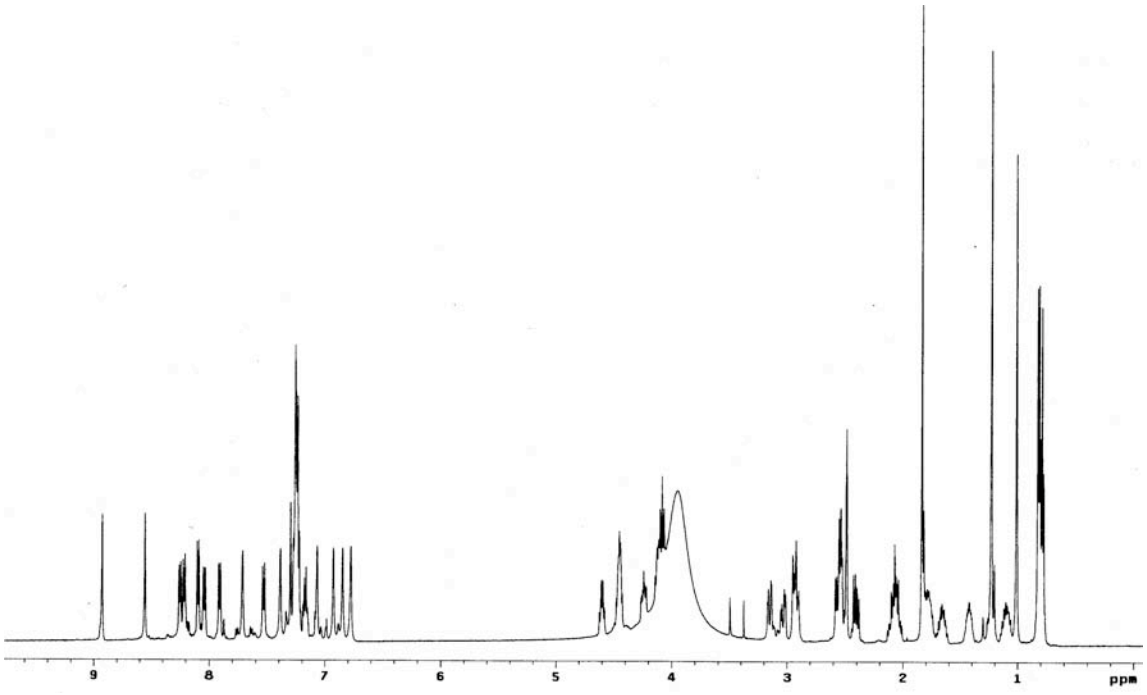
80-ms TOCSY of peptidomimetic **5** (90% H_2O/10% D_2O)



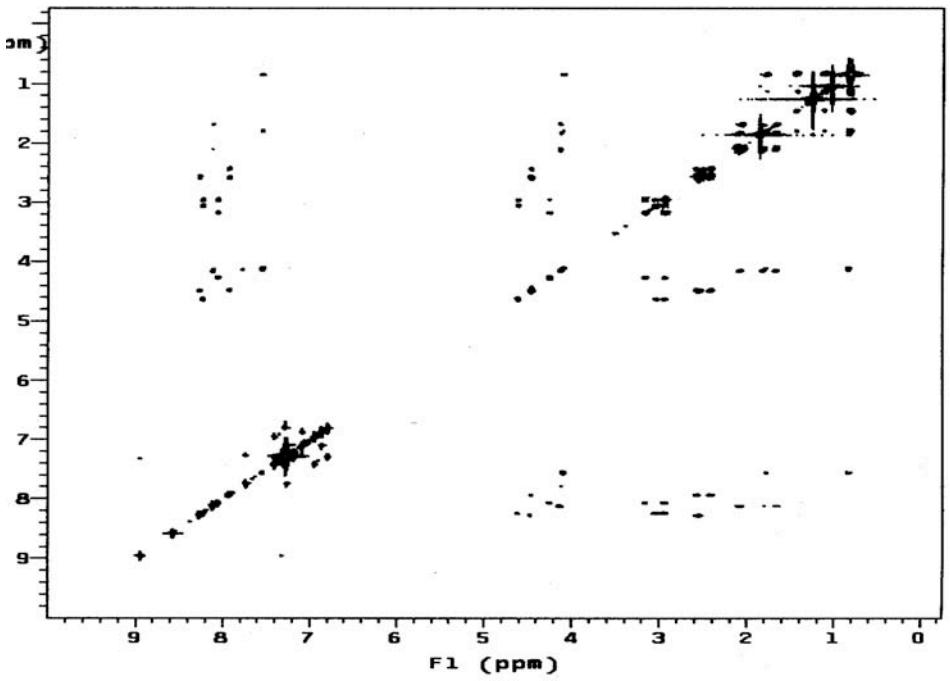
^{1}H NMR of peptidomimetic **6** (90% H_2O/10% D_2O)



80-ms TOCSY of peptidomimetic **6** (85% H_2O/15% $(CF_3)_2CDOD$)



^{1}H NMR of peptidomimetic **7** (DMSO-d6)



80-ms TOCSY of peptidomimetic **7** (DMSO-d6)

Table E8. ROEs of Peptidomimetic 1 in 80%H_2O:20%$(CF_3)_2CDOD$

proton 1	proton 2	ROE	proton 1	proton 2	ROE
Phe2NH	Tyr3NH	strong	Tyr3NH	Aib4NH	strong
Aib4NH	Ile5NH	strong	Ile5NH	Leu6NH	strong
Leu6NH	His7NH	strong	His7NH	Acp8NH	strong
Gln9NH	His10NH	strong	His10NH	Asn11NH	strong
Asn11NH	Aib12NH	strong	Aib12NH	Phe13NH	strong
Phe13NH	Ile14NH	strong	Ile14NH	Asn15NH	strong
Gln9□H	His10NH	strong	Asp1□H	Asp1NH	weak
His10□H	His10NH	medium	Ile14□H	Asn15NH	strong
Asn15□H	Asn15NH	weak	Ile5□H	Leu6NH	weak
Ile5□H	Acp8NH	weak	Ile5□H	Ile5NH	medium
Tyr3□H	His7NH	weak	Tyr3□H	Leu6NH	weak
Tyr3□H	Tyr3NH	medium	Leu6□H	His7NH	weak
Leu6□H	Leu6NH	weak	Ile14□H	Ile14NH	medium
Gln9□H	Gln9NH	medium	Phe2□H	Phe2NH	medium
Phe2□H	Tyr3NH	weak	Phe2□H	Ile5NH	weak
His7□H	His7NH	medium	His7□H	Acp8NH	weak
Phe13□H	Ile14NH	medium	Phe13□H	Phe13NH	medium
Asp1□H	Aib4NH	medium	Asn11□H	Asn11NH	medium
Asn11□H	Aib12NH	medium	Asn11□H	Ile14NH	medium
His10□H	Phe13NH	weak	Tyr3□H	(Tyr3)2,6H	strong
Phe2□H	(Phe2)2,6H	strong	Phe13□H	(Phe13)2,6H	strong
Tyr3□H	(Tyr3)2,6H	strong	Phe2□H	(Phe2)2,6H	strong
Phe13□H	(Phe13)2,6H	strong	Asp1□H	Aib4□H	weak
Phe2□H	Ile5□H	medium	Tyr3□H	Leu6□H	weak

Table E9. NOEs of Peptidomimetic 4 in 85%H_2O/15%$(CF_3)_2CDOD$

proton 1	proton 2	ROE	proton 1	proton 2	ROE
Asp1NH	Gln2NH	strong	Gln2NH	Gln3NH	strong
Gln3NH	Asn4NH	strong	Tyr7NH	Aib8NH	strong
Aib8NH	Ala9NH	strong	Ala9NH	Leu10NH	strong
Ala13NH	Arg14NH	medium	Aib17NH	Ile18NH	strong
Ile18NH	Gln19NH	strong	Gln19NH	Asn20NH	strong
Asn20NH	(amide)NH	weak	Asp1□H	Gln2NH	medium
Gln2□H	Gln2NH	strong	Tyr7□H	Aib8NH	medium
Ala5□H	Aib8NH	medium	Phe6□H	Aib8NH	weak
Asn4□H	Aib8NH	weak	Asp1□H	Asp1NH	weak
Gln3□H	Gln3NH	strong	Gln2□H	Gln3NH	strong
Ala9□H	Ala9NH	medium	Tyr7□H	Tyr7NH	medium
Ile18□H	Gln19NH	medium	Arg14□H	Arg14NH	medium
Gln19□H	Asn20NH	medium	Ala13□H	Arg14NH	medium
Gln19□H	Gln19NH	medium	Gln3□H	Phe6NH	medium
Arg14□H	Asn15NH	medium	Ala13□H	Ala16NH	weak
Leu10□H	Leu10NH	medium	Leu10□H	His11NH	medium
Ala9□H	Leu10NH	medium	Ala9□H	Acp12NH	weak
Ala16□H	Ala16NH	weak	Ala16□H	Aib17NH	medium
Gln2□H	Ala5NH	medium	Ala5□H	Ala5NH	medium
Ala16□H	Gln19NH	medium	Ala13□H	Ala13NH	weak
Phe6□H	Ala9NH	weak	Phe6□H	Tyr7NH	weak
Phe6□H	Phe6NH	medium	His11□H	His11NH	medium
Asn4□H	Tyr7NH	medium	Asn4□H	Asn4NH	medium
Asn4□H	Ala5NH	medium	Asn15□H	Asn15NH	medium
Asn15□H	Ala16NH	medium	Asn20□H	Asn20NH	weak
Asp1□H	Asn4NH	weak	Ile18□H	Ile18NH	strong
Ala16□H	Ile18NH	weak	Tyr7□H	(Tyr7)2,6H	strong
Phe6□H	(Phe6)2,6H	medium	Asp1□H	Asn4□H	weak
Gln2□H	Ala5□H	medium	Gln3□H	Phe6□H	weak
Asn4□H	Tyr7□H	weak	Ala5□H	Aib8□H	medium
Phe6□H	Ala9□H	medium	Tyr7□H	Leu10□H	medium
Arg14□H	Aib17□H	medium	Ala16□H	Gln19□H	medium
Phe6□H	Aib8NH	weak	Ala16□H	Ile18NH	weak

Table E10. NOEs of Peptidomimetic 5 in 85%H_2O/15%$(CF_3)_2CDOD$

proton 1	proton 2	ROE	proton 1	proton 2	ROE
Asp1NH	Ala2NH	medium	Leu18NH	Lys19NH	strong
Aib5NH	Ala6NH	strong	Lys19NH	Asn20NH	strong
Ile15NH	Gln16NH	strong	Tyr4NH	Aib5NH	strong
Ala13NH	Aib14NH	strong	Aib14NH	Ile15NH	strong
Arg11NH	Asn12NH	medium	Ala6NH	Leu7NH	strong
Leu7NH	His8NH	strong	Asn12NH	Ala13NH	medium
Phe3NH	Tyr4NH	strong	Ala2NH	Phe3NH	strong
Asn20NH	(amide)NH_2	medium	Tyr4NH	(Tyr4)2,6H	medium
(Tye4)2,6H	Aib5NH	medium	Ala10NH	Arg11NH	strong
Gln16NH	Ser17NH	strong	Asp1□H	Asp1NH	weak
Asp1□H	Ala2NH	weak	Ala13□H	Aib14NH	medium
Leu18□H	Lys19NH	weak	Lys19□H	Lys19NH	medium
Ala2□H	Ala2NH	weak	Ala10□H	Arg11NH	medium
Tyr4□H	Aib5NH	weak	Ala2□H	Aib5NH	weak
Ser17□H	Ser17NH	medium	Leu18□H	Leu18NH	medium
Ser17□H	Ser17NH	medium	Ile15□H	Leu18NH	weak
Gln16□H	Gln16NH	medium	Ala6□H	Ala6NH	medium
Arg11□H	Asn12NH	medium	Phe3□H	Ala6NH	weak
His8□H	His8NH	medium	Asn12□H	Asn12NH	medium
Asn20□H	Asn20NH	medium	Ile15□H	Ile15NH	strong
Tyr4□H	Tyr4NH	medium	Leu7□H	Leu7NH	medium
Ala10□H	Ala10NH	medium	Ala13□H	Ala13NH	weak
Phe3□H	Tyr4NH	weak	Asn12□H	Ala13NH	weak
Asp1□H	Tyr4NH	weak	Ala2□H	Phe3NH	weak
Phe3□H	Phe3NH	medium	Phe3□H	(Phe3)2,6H	medium
His8□H	(His8)4H	weak	Tyr4□H	(Tyr4)2,6H	medium
Tyr4NH	Leu7□H	medium	Tyr4□H	Aib5NH	medium
Asp1NH	Asp1□H	weak	Gln16□H	Gln16NH	medium
His8□H	His8NH	strong	Asn12□H	Asn12NH	strong
Asn20□H	Asn20NH	strong	Tyr4□H	Tyr4NH	strong
Asn12□H	Ala13NH	weak	Phe3□H	Phe13NH	strong
Asn20□H	(amide)NH_2	weak	Phe3□H	(Phe3)2,6H	strong
Tyr4□H	(Tyr4)2,6H	strong	His8□H	(His8)4H	weak
Gln16□H	Ser17NH	strong	Gln16□H	Gln16NH	strong
Aib5□H	Aib5NH	strong	Lys19□H	Lys19NH	strong
Arg11□H	Arg11NH	medium	Leu18□H	Leu18NH	medium
Lys19□H	Asn20NH	weak	Ile15□H	Ile15NH	strong
Ala10□H	Aib14NH	strong	Ala2□H	Ala2NH	strong

Table E10. (Continued)

proton 1	proton 2	ROE	proton 1	proton 2	ROE
Leu7αH	Arg11NH	strong	Aib14αH	Ser17NH	medium
Leu18αH	Leu18NH	medium	Ala6αH	Ala6NH	strong
Aib5αH	Ala6NH	strong	Leu7αH	His8NH	strong
Asn12NH	Gln16αH	medium	Aib14NH	Leu18αH	medium
Tyr4NH	Aib5αH	weak	Ile15NH	Lys19αH	medium
Ile15NH	Lys19αH	strong	Leu7NH	Leu7αH	strong
Leu7NH	Leu7αH	weak	Leu18NH	Leu18αH	weak
Ala2αH	Phe3NH	strong			

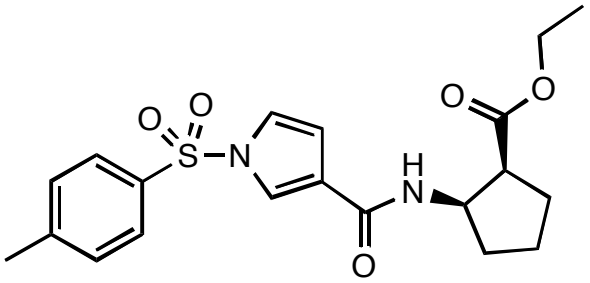
Table E11. ROEs of Peptidomimetic 6 in 85%H_2O/15%$(CF_3)_2CDOD$

proton 1	proton 2	ROE	proton 1	proton 2	ROE
AspNH	PheNH	strong	TyrNH	AibNH	strong
TyrNH	PheNH	strong	LeuNH	HisNH	strong
IleNH	AibNH	strong	IleNH	LeuNH	strong
HisαH	HisNH	weak	AspαH	AspNH	weak
LeuαH	HisNH	weak	LeuαH	LeuNH	weak
PheαH	PheNH	weak	TyrαH	TyrNH	weak
PheαH	IleNH	weak	AspαH	IleNH	weak
IleαH	IleNH	weak	PheNH	Phe2 or 6H	medium
TyrαH	Tyr2 or 6H	medium	TyrαH	Tyr2 or 6H	medium
PheαH	Phe2 or 6H	medium	TyrαH	TyrNH	strong
PheαH	PheNH	strong	HisαH	HisNH	medium
LeuαH	LeuNH	medium	IleNH	AibαH	strong
IleNH	IleαH	weak	IleNH	IleαH	medium
AibαH	Tyr2, or 6H	strong	TyrαH	LeuNH	medium
PheαH	Phe2 or 6H	strong	AibαH	AibNH	strong
AspαH	AibαH	weak	PheαH	IleαH	medium
TyrαH	LeuαH	medium			

APPENDIX F
EXPERIMENTAL FOR CHAPTER IV

General Procedures. -Amino acids used were of the L-configuration. All chemicals were obtained from commercial suppliers and directly used without further purification. TentaGel S PHB Resin was obtained from Rapp Polymere and amino acids were obtained from Advanced Chem Tech. 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), *N*- hydroxybenzotriazole (HOBt), di-iso-propylethylamine, trifluoroacetic acid (TFA), CH_2Cl_2, DMF, 4-nitrophenylchloroformate, N-methylmorpholine, Poly(propylene imine)dendrimer (DAB-Am-n), propargylamine, lithium hydroxide monohydrate, triphosgene, bromoacetyl bromide, substituted aniline, sodium azide, tris(hydroxymethyl)aminomethane (Tris), 1,3-dimethoxybenzene and copper(I) iodide were purchased from Aldrich. Epoxy-activated sepharose 6B was purchased from Amersham biosciences. DMF was stored over 4Å molecular sieves for a few days before use. For solid phase synthesis, fritted polypropylene syringes were purchased from Torviq, and a manually controlled shaking apparatus was used. Reverse phase high performance liquid chromatography (RP-HPLC) was carried out using BACKMAN system on Vydac C-18 columns of the following dimensions: 25 x 2.2 cm for preparation, and 25 x 0.46 cm for analysis. All HPLC experiments were performed using gradient conditions. The eluants used were solvent A (H_2O with 0.1% TFA) and solvent B (CH_3CN with 0.1% TFA). Flow rates used were 10 mL/min for preparative HPLC, and 1.0 mL/min for analytical HPLC. All NMR spectra were recorded on Varian or Inova instruments at 300 MHz for ^1H and 75 MHz for ^{13}C; or 500 MHz for ^1H and 125 MHz for ^{13}C. NMR chemical shifts are expressed in ppm relative to solvent ($CDCl_3$: 7.27 ppm for ^1H and 77 ppm for ^{13}C, acetone-d6: 2.04 ppm for ^1H and 29.9 ppm for ^{13}C; DMSO-d6: 2.50 ppm for ^1H and 39.5 ppm for ^{13}C; CD_3CN: 1.93 ppm for ^1H and 1.3 ppm for ^{13}C), and coupling constants were measured in Hz. Multiplicities in ^1H NMR were

reported as s (singlet), d (doublet), t (triplet), m (multiplet). FT-IR spectra were obtained using the solution of product in CH_2Cl_2 and 4021 GALAXY series instrument. MALDI-TOF and ESI mass spectra were obtained from the Mass Spectrometry Applications Laboratory at Texas A&M University. Thin layer chromatography was carried out using silica gel 60 F254 plates. Flash column was performed using silica gel (230-600 mesh).



11

Procedure

To a suspension of *N*-tosyl-3-pyrrolecarboxylic acid (570 mg, 2.15 mmol) in 30 ml dichloromethane was added ethylimethylaminopropyl carbodiimide (453 mg, 2.37 mmol) and 4-(dimethylamino)pyridine (867 mg, 7.1 mmol) at 0˚C. After 20 min, ethyl-cis-2-amino-1-cyclopentanecarboxylate hydrochloride (Eacpc) (500 mg, 2.58 mmol) dissolved in 6 ml dichloromethane was then added. The mixture was stirred for 18 h at 25 ˚C and concentrated *in vacuo*. The residue was added 1M HCl and then extracted with ethyl acetate (20 mL X 4). The combined extracts were washed with saturated sodium hydrogencarbonate and saturated sodium chloride, dried over anhydrous magnesium sulfate. The solvent was removed *in vacuo*. The crude material was chromatographed with ethyl acetate/hexane (30-50%) to afford product **11** as white solid (800 mg, 92%).

$\boldsymbol{R_f}$ = 0.47 (1:1 ethyl acetate/hexane)

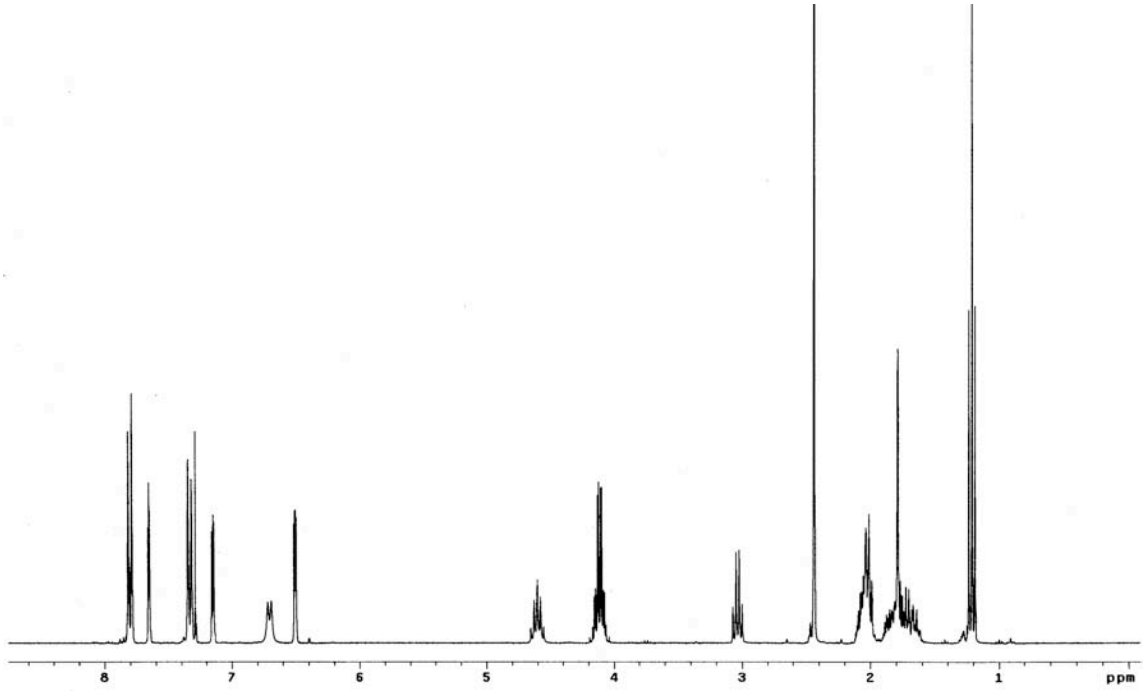
^{1}H NMR (300 MHz, $CDCl_3$)

= 7.8 (d, *J*=8.1, 2H), 7.66 (t, J=2.01, 1H), 7.33 (d, J=8.1, 2H), 7.15 (q, J=3.3, 1H), 6.71 (d, J=8.1, 1H), 6.51 (q, J=1.68, 1H), 4.61 (m, J=7.5, 1H), 4.03-4.20 (m, 2H), 3.04 (q, J=7.5, 1H), 2.44 (s, 3H), 1.95-2.13 (m, 3H), 1.59-1.92 (m, 3H), 1.21 (t, J=7.5, 3H)

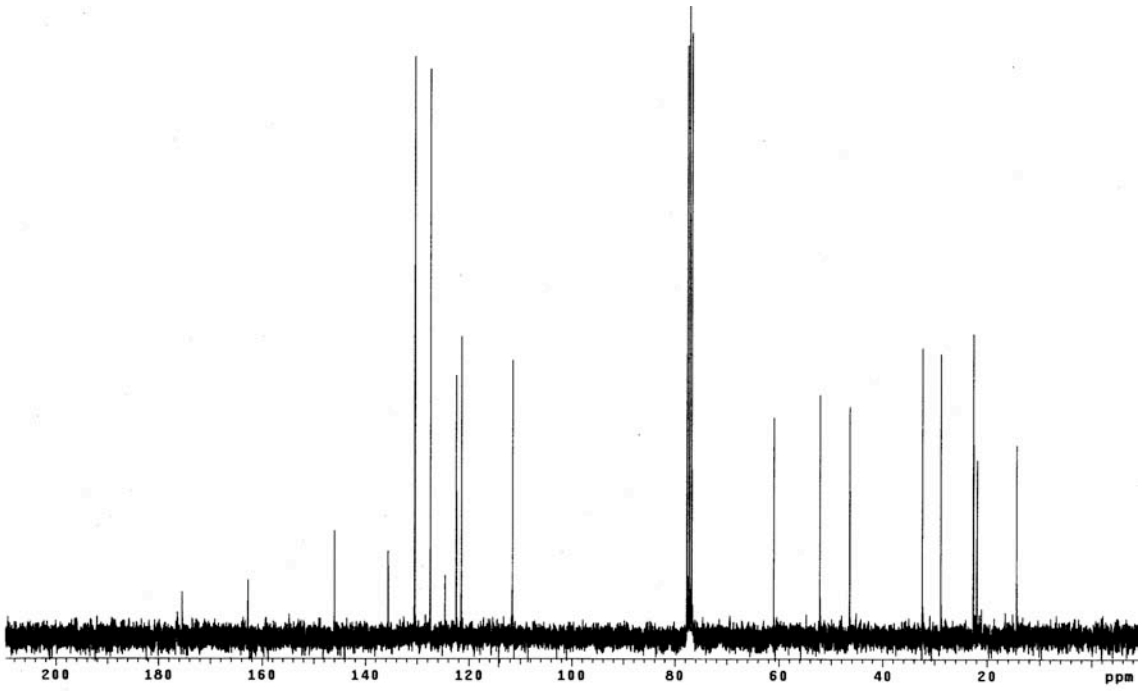
^{13}C NMR (75 MHz, $CDCl_3$)

= 175.4, 162.7, 146.0, 136.0, 130.5, 127.5, 124.6, 122.5, 121.5, 111.5, 61.0, 52.2, 46.4, 32.4, 28.9, 22.6, 21.9, 14.4

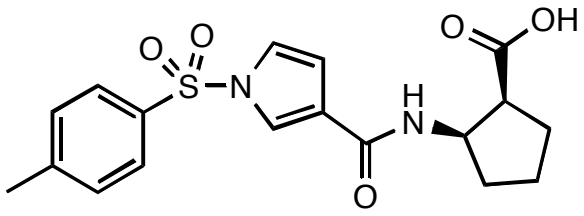
MS (ESI) calc'd for $C_{20}H_{25}N_2O_5S$ $(M+H)^+$ 405.14, found 405.20



^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)



15

Procedure

Lithium hydroxide monohydrate (40.5 mg, 0.965 mmol) was added to a solution of **11** (300 mg, 0.742mmol) in THF/methanol/H2O (3:1:1) (2.5 ml). The mixture was stirred at 0˚C for 8 h. The solvent was removed *in vacuo*, the residue was

diluted with 3 ml of water and then acidified with 2M HCl (0.5 ml) up to a PH 2-3. The resulting solution was extracted with 3 x 10 ml ethyl acetate, dried over anhydrous sodium sulfate. The dried solution was concentrated *in vacuo* to give the crude material which was recrystallized with 50% ethanol in water to afford product **15** as white solid (237 mg, 85%).

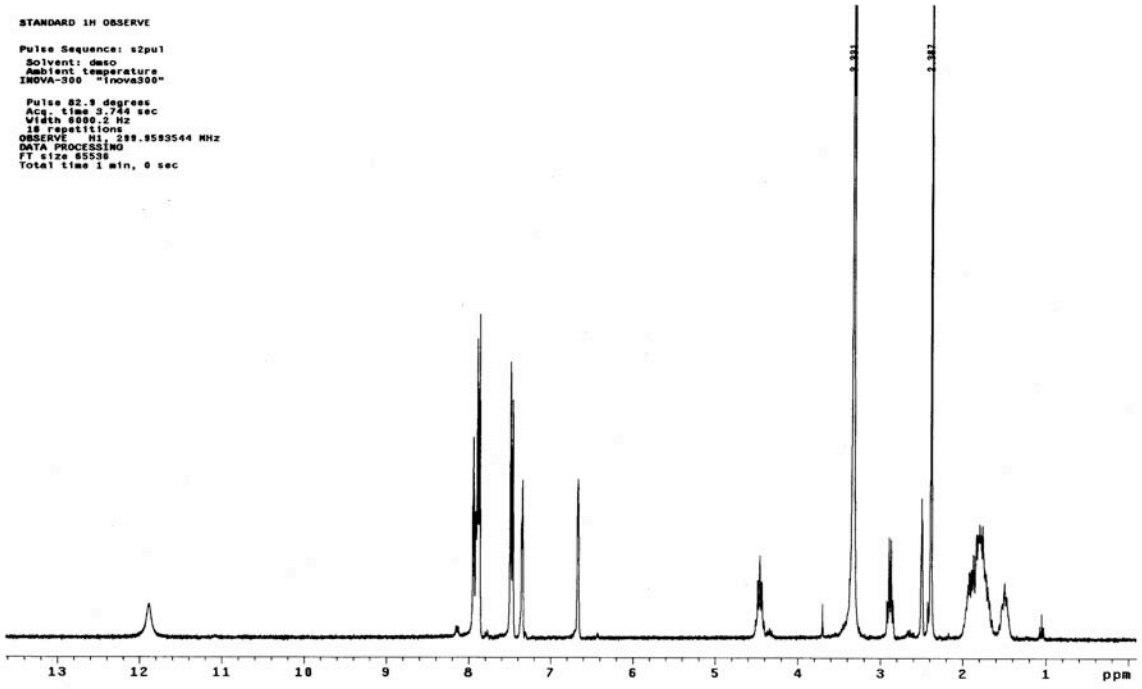
^{1}H NMR (300 MHz, DMSO-d6**)**

= 11.89 (bs, 1H), 7.95 (m, 1H), 7.91 (d, J=3.9, 1H), 7.88 (d, J=7.8, 2H), 7.48 (d, J=7.8, 2H), 7.35 (m, 1H), 6.67 (m, 1H), 4.46 (m, J=7.2, 1H), 2.88 (q, J=7.2, 1H), 2.39 (s, 3H), 1.62-2.01 (m, 5H), 1.42-1.58 (m, 1H)

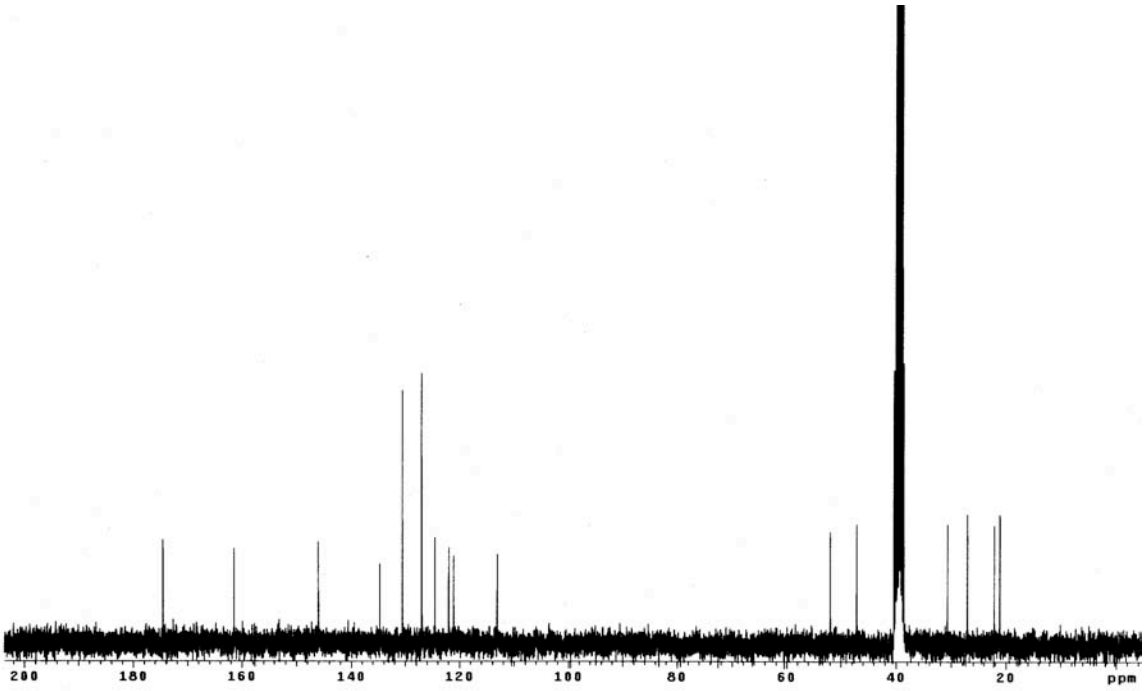
^{13}C NMR (75 MHz, DMSO-d6)

= 174.5, 161.4, 146.0, 134.7, 130.5, 127.0, 124.5, 121.9, 121.1, 113.1, 52.0, 47.1, 30.7, 27.1, 22.1, 21.1

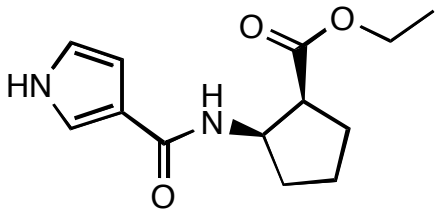
MS (ESI) $(M+H)^{+}$ calc'd 377.11, found 377.11



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



12

Procedure

To a suspension of pyrrole-3-carboxylic acid (717 mg, 6.45 mmol) in 20 ml N, N-dimethylformamide and 30 ml dichloromethane was added 1,3-diisopropyl carbodiimide (1212 uL, 7.74 mmol) and 4-(dimethylamino)pyridine (1735 mg, 14.2 mmol) at 0 ˚C. After 20 min, ethyl-cis-2-amino-1-cyclopentanecarboxylate

hydrochloride (1.5 g, 7.74 mmol) dissolved in 30 ml dichloromethane was then added. The mixture was stirred for 18 h at 25 ˚C, and concentrated *in vacuo*. The residue was added 1M HCl and then extracted with with ethyl acetate (20 mL X 4). The combined extracts were washed with saturated sodium hydrogencarbonate and saturated sodium chloride, dried over anhydrous magnesium sulfate. The solvent was removed *in vacuo*. The crude material was chromatographed with ethyl acetate/hexane (60%) to afford product **12** as a slightly brown solid (1.359 g, 84%).

R_f = 0.31 (3:1 ethyl acetate/hexane)

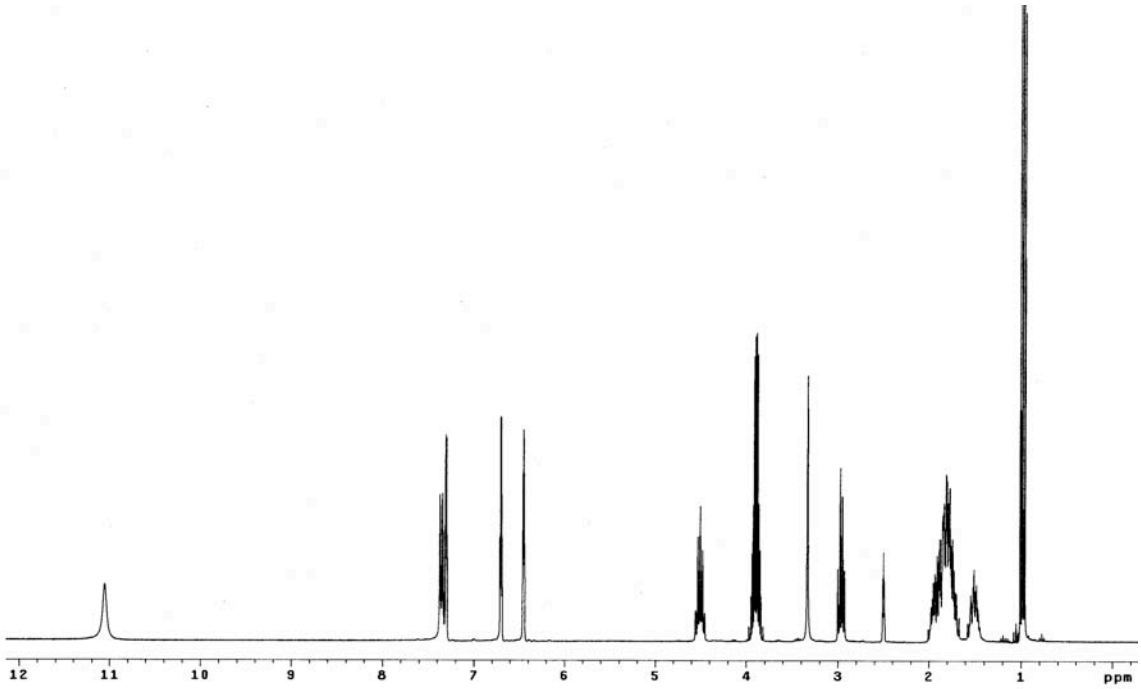
^{1}H NMR (300 MHz, DMSO-d6)

 = 11.05 (s, 1H), 7.36 (d, J=8.1, 1H), 7.30 (m, 1H), 6.70 (q, J=2.28, 1H), 6.45 (m, 1H), 4.50 (m, J=7.2, 1H), 3.81-3.98 (m, 2H), 2.97 (q, J=7.2, 1H), 1.66-2.01 (m, 5H), 1.43-1.59 (m, 1H), 0.99 (t, J=7.2, 3H)

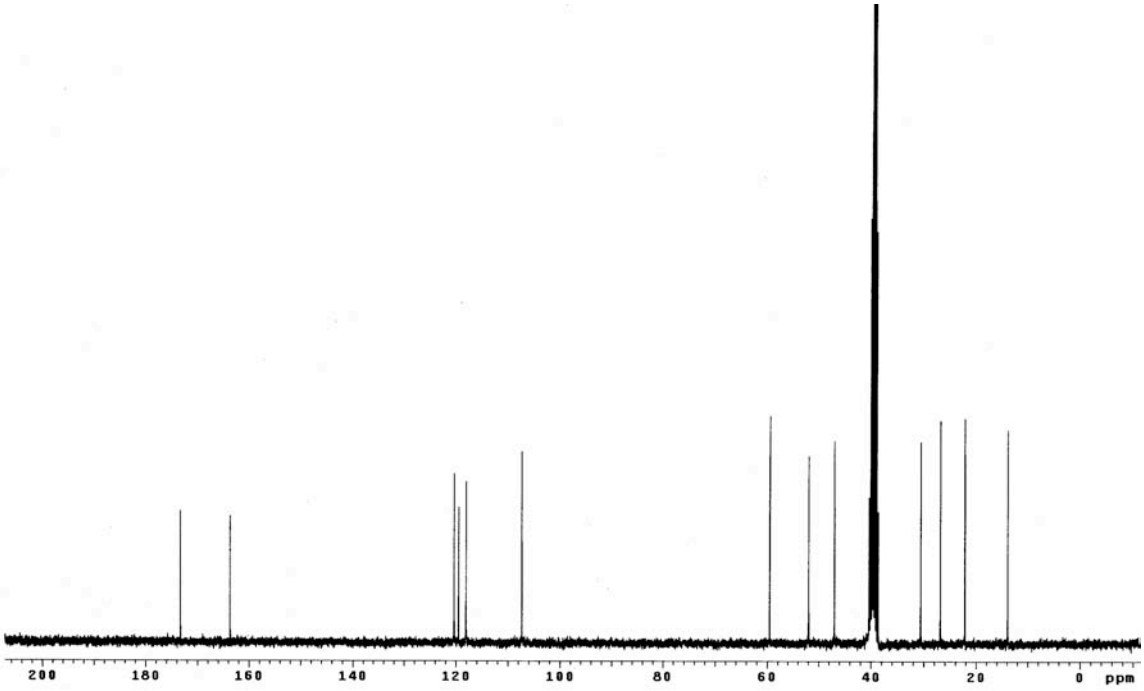
^{13}C NMR (75 MHz, DMSO-d6)

 = 173.3, 163.7, 120.4, 119.5, 118.1, 107.3, 59.5, 52.0, 47.0, 30.6, 26.9, 22.2, 13.9

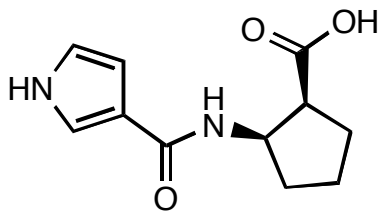
MS (ESI) $(M+H)^+$ calcd 251.13, found 251.14



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



22

Procedure

Lithium hydroxide monohydrate (231 mg, 5.505 mmol) was added to a solution of compound **12** (1.06 mg, 4.23 mmol) in THF/MeOH/H_2O (3:2:1) (9 ml). The mixture was stirred at 25 ˚C for 1 h. The solvent was removed *in vacuo*, the residue was diluted with 5 ml of water and then acidified with 2M HCl (6 ml) up to a PH 2-3. The resulting solution was extracted with 3 x 50 ml ethyl acetate, dried over anhydrous sodium sulfate. The dried solution was concentrated *in vacuo* to give the crude material which was recrystallized with 50% ethanol in water to afford product **22** as white solid (828 mg, 90%).

R_f = 0.27 (3:1:0.1 ethyl acetate/hexane/acetic acid)

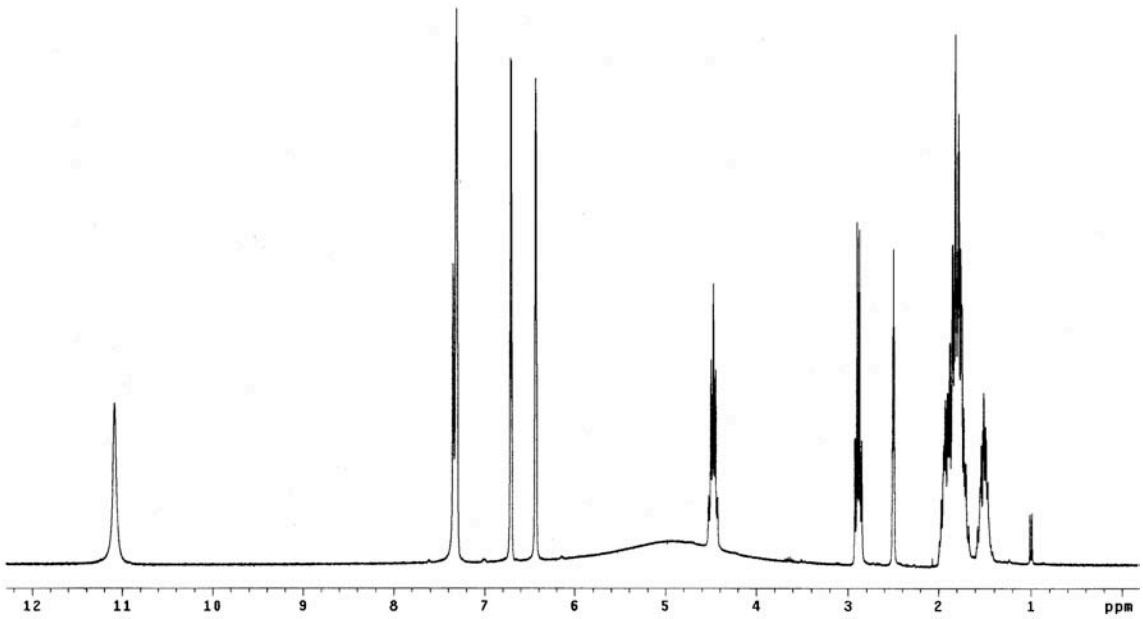
^{1}H NMR (300 MHz, DMSO-d6)

= 11.10 (s, 1H), 7.28-7.37 (m, 2H), 6.71 (q, J=2.29, 1H), 6.44 (m, 1H), 4.92 (bs, 1H), 4.48 (m, J=7.2, 1H), 2.89 (q, J=7.2, 1H), 1.66-2.01 (m, 5H), 1.42-1.59 (m, 1H)

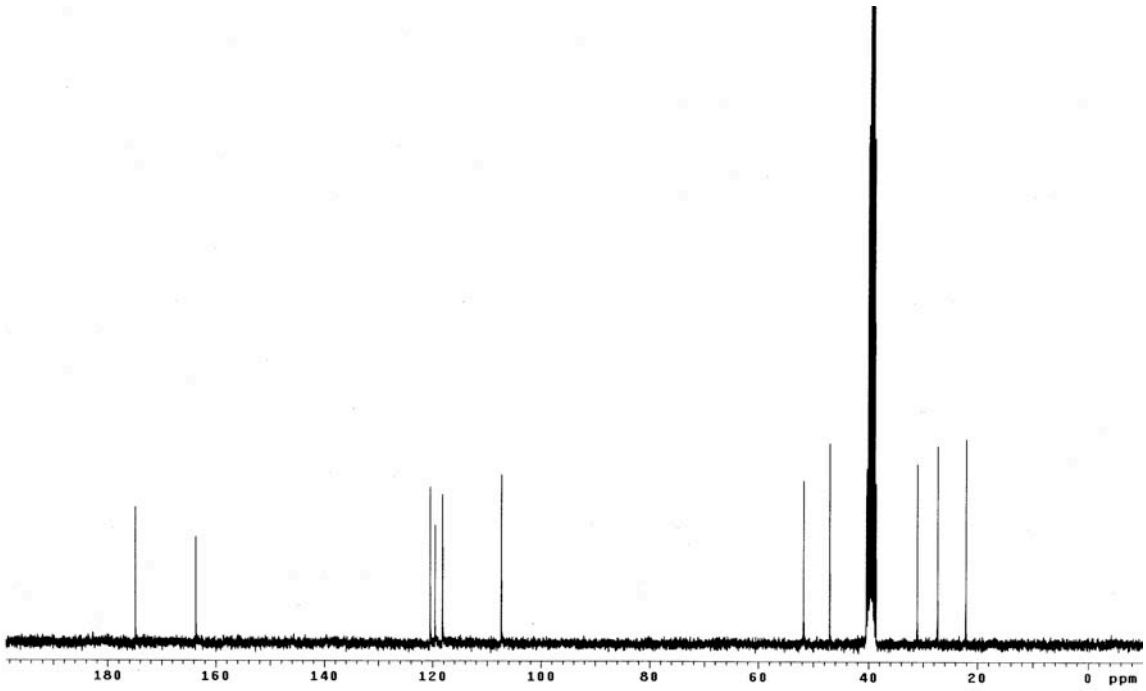
^{13}C NMR (75 MHz, DMSO-d6)

= 174.9, 163.6, 120.5, 119.6, 118.2, 107.4, 51.7, 47.0, 31.1, 27.3, 22.1

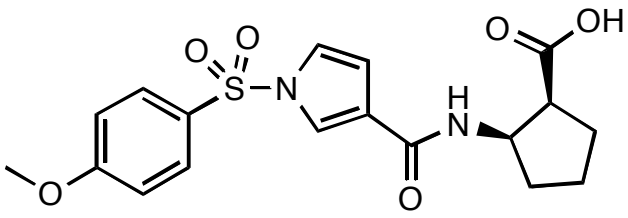
MS (ESI) $(M-H)^+$ calc'd 221.10, found 221.10



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



16

Procedure

Intermediate **22** (71.4 mg, 0.32 mmol) **w**as dissolved in 10 ml dichloroethane and 3 ml *N*, *N*-dimethylformamide, and cooled to 0 ˚C. Ground sodium hydroxide (76.8 mg, 1.92 mmol) was added to the solution and stirred for 1 h. A solution of 4-methoxybenzenesulfonyl chloride (198.7 mg, 0.96 mmol) in 6 ml dichloroethane was added dropwise. After 40 min the mixture was allowed to 25 ˚C and stirred 14 h. The reaction was quenched by addition of 10 ml 1M HCl. The organic layer was separated, and the aqueous layer was extracted with dichloromethane (3 x 20ml). The combined organic extracts were washed with saturated sodium chloride, dried over sodium sulfate. Removal of the solvent *in vacuo* gave the crude material which was chromatographed with methanol/ethyl acetate (0-20%) to afford product **16** as white solid (126 mg, 76%).

R_f = 0.51 (3:1:0.1 ethyl acetate/hexane/acetic acid)

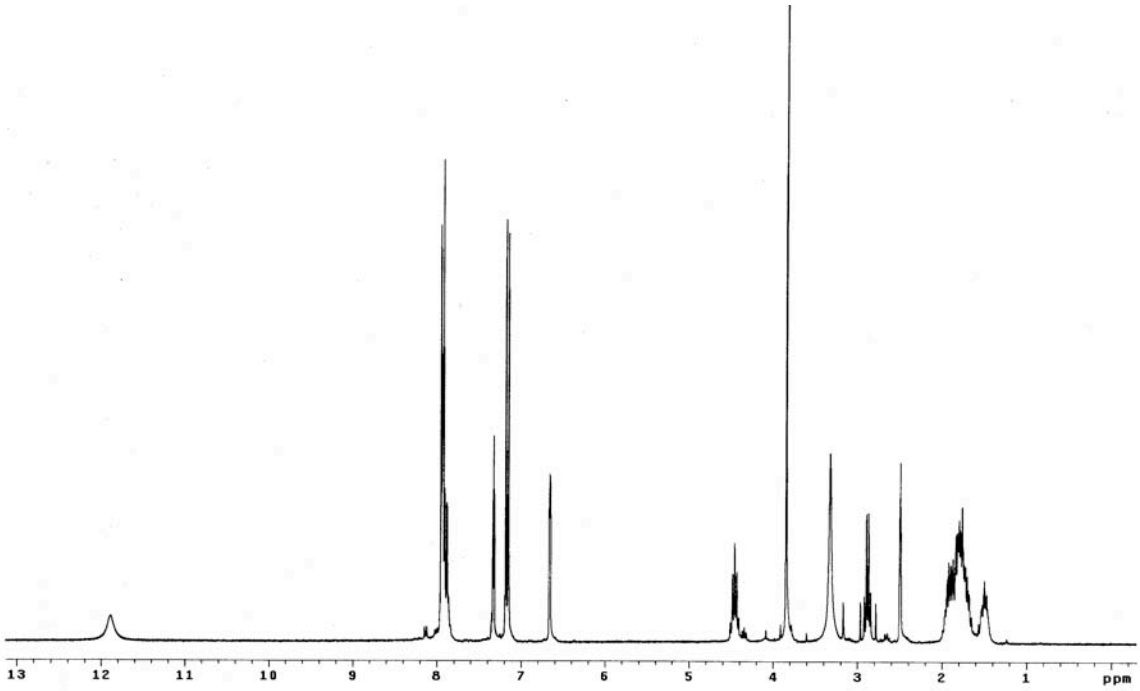
^{1}H NMR (300 MHz, DMSO-d6**)**

= 11.89 (bs, 1H), 7.92-7.96 (m, 3H), 7.89 (d, J=8.1, 1H), 7.33 (m, 1H), 7.17 (d, J=8.4, 2H), 6.66 (m, 1H), 4.46 (m, J=7.5, 1H), 3.85 (s, 3H), 2.88 (q, J=7.7, 1H), 1.62-2.01 (m, 5H), 1.42-1.58 (m, 1H)

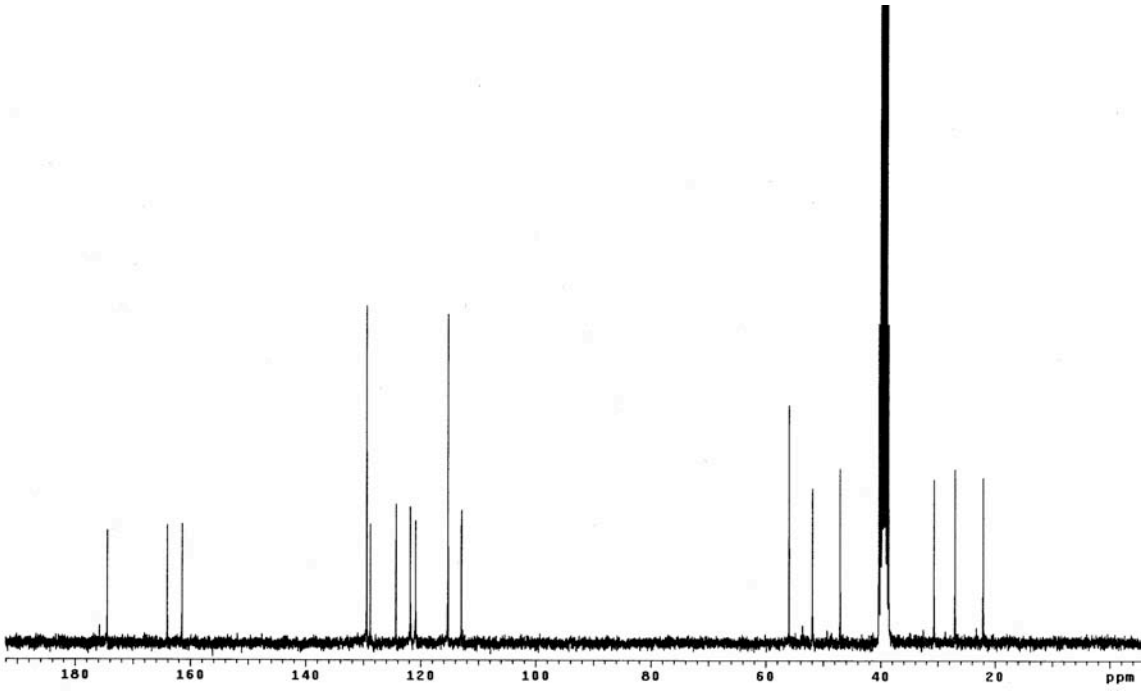
^{13}C NMR (75 MHz, DMSO-d6)

= 174.5, 164.0, 161.5, 129.5, 128.8, 124.4, 121.8, 120.9, 115.3, 112.9, 56.0, 51.9, 47.1, 30.7, 27.1, 22.1

MS (ESI) $(M+H)^+$ calc'd 393.10, found 392.9



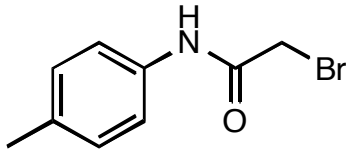
^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

Syntheses of 2-Bromo-*N*-Aryl Acetamides 23

A solution of aryl amine (0.08 mol) in CH_2Cl_2 (120 mL) and triethylamine (12 mL, 0.08 mol) was cooled in an ice bath with stirring, and a solution of bromoacetyl bromide (6.97 mL, 0.08 mol) in CH_2Cl_2 (50 mL) was added dropwise. After the completion of addition the mixture was stirred about 1 h at 0 ˚C, then stirred 4 h at 25 ˚C. To this reaction mixture was added 130 ml CH_2Cl_2. The mixture was washed successively with 1M HCl (2 x 30 mL), $NaHCO_3$ (1 x 30 mL) and H_2O (1 x 30 mL), and dried over anhydrous Na_2SO_4. Finally the mixture was filtered and evaporated *in vacuo*. The crude material was recrystallized with 60 mL CH_3CH_2OH and gave 2-bromo-*N*-Aryl acetamides in yields from 86 to 94%. All compounds gave satisfactory MS data and were verified by their ^{1}H NMR and ^{13}C NMR. The spectral data of three compounds: **23u**, **23z** and **23v**, were taken for illustration.



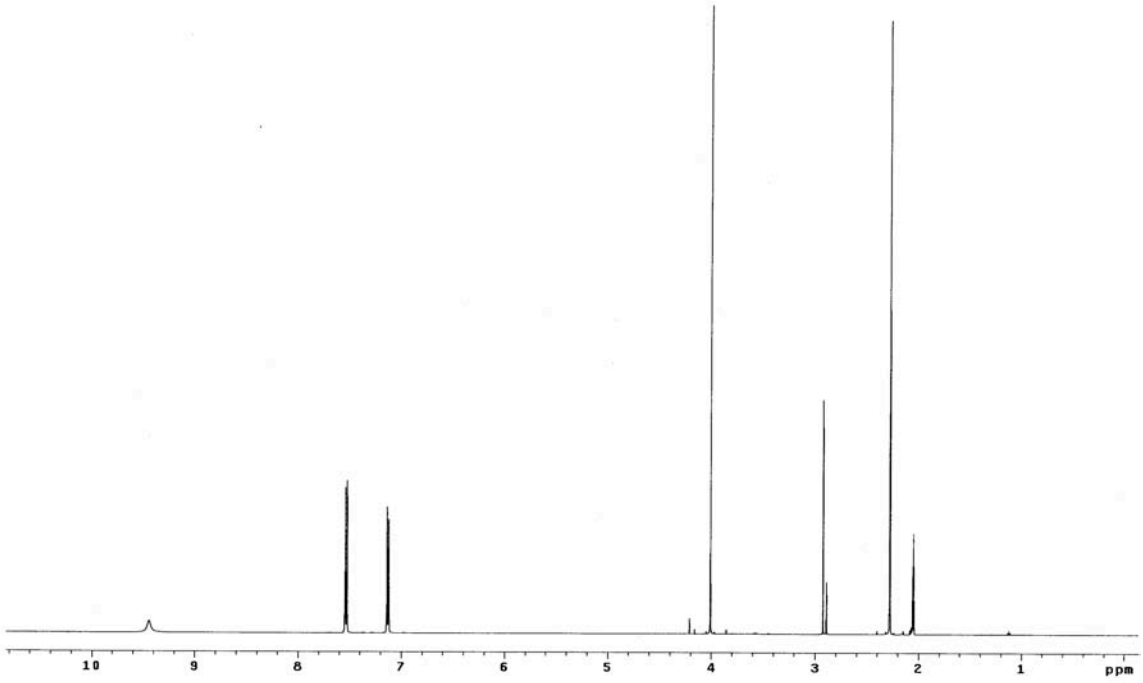
23u

^{1}H NMR (500 MHz, acetone-d6)

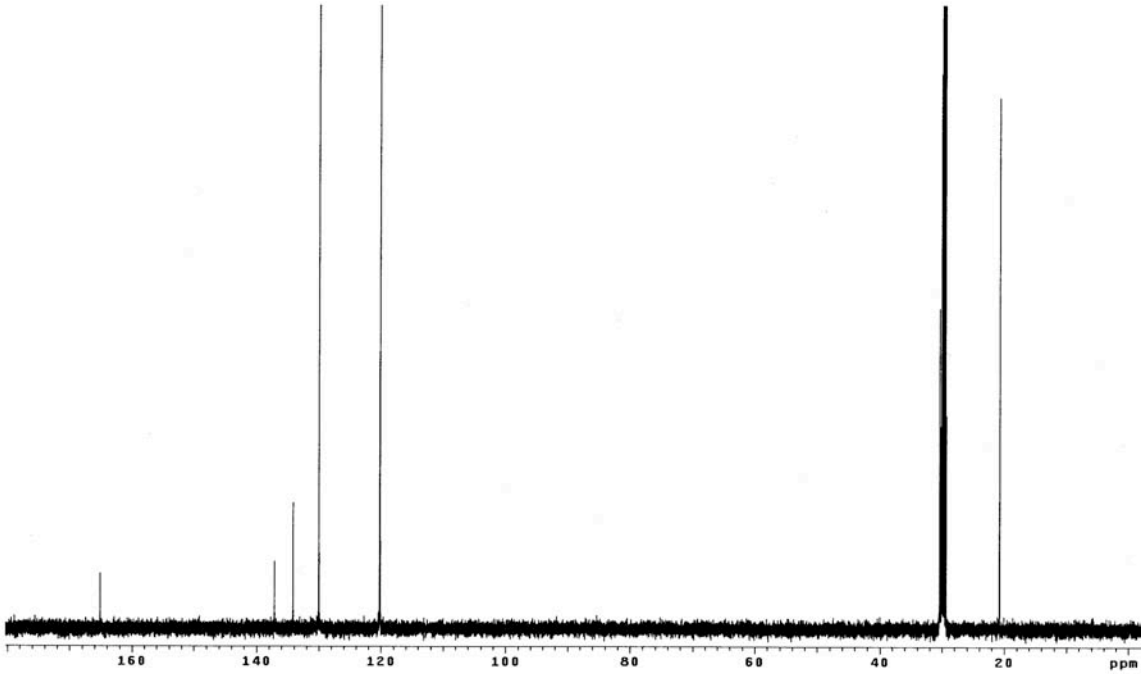
= 9.44 (bs, 1H), 7.53 (d, J=8.0, 2H), 7.13 (d, J=8.0, 2H), 4.01 (s, 2H), 2.28 (s, 3H)

^{13}C NMR (125 MHz, acetone-d6)

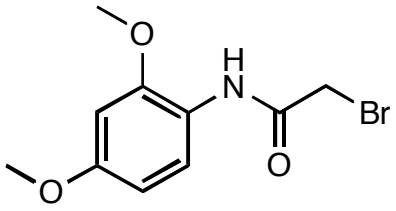
= 165.1, 137.2, 134.2, 130, 120.2, 30.4, 20.8



^{1}H NMR (acetone-d6)



^{13}C NMR (acetone-d6)



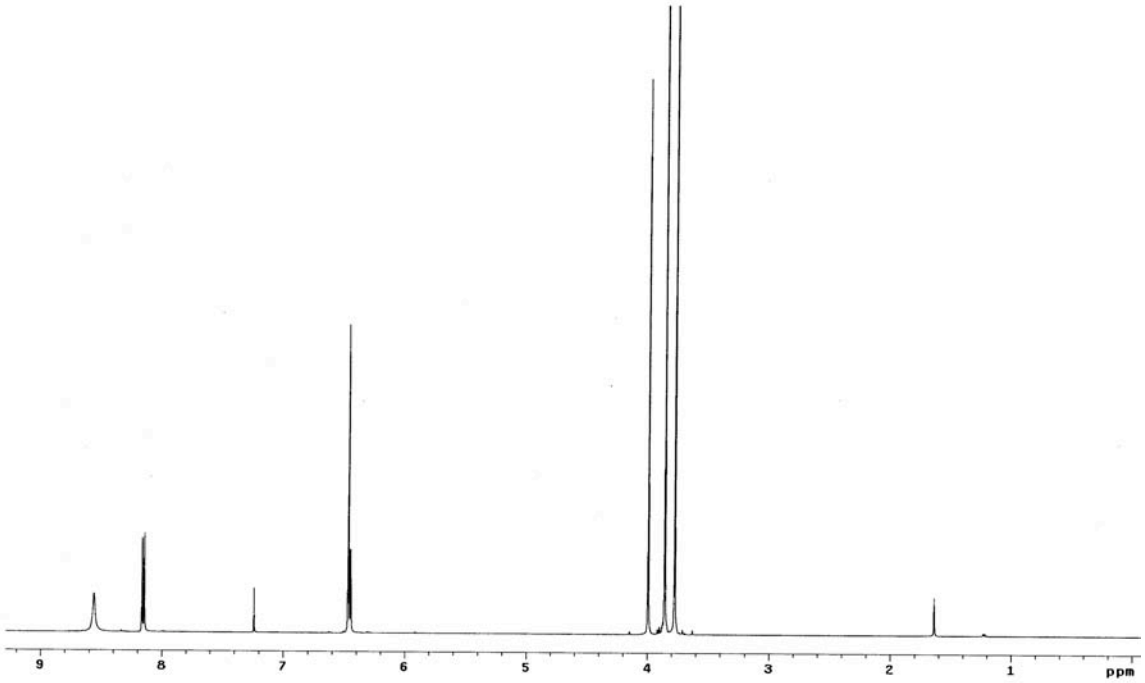
23z

^{1}H NMR (500 MHz, $CDCl_3$)

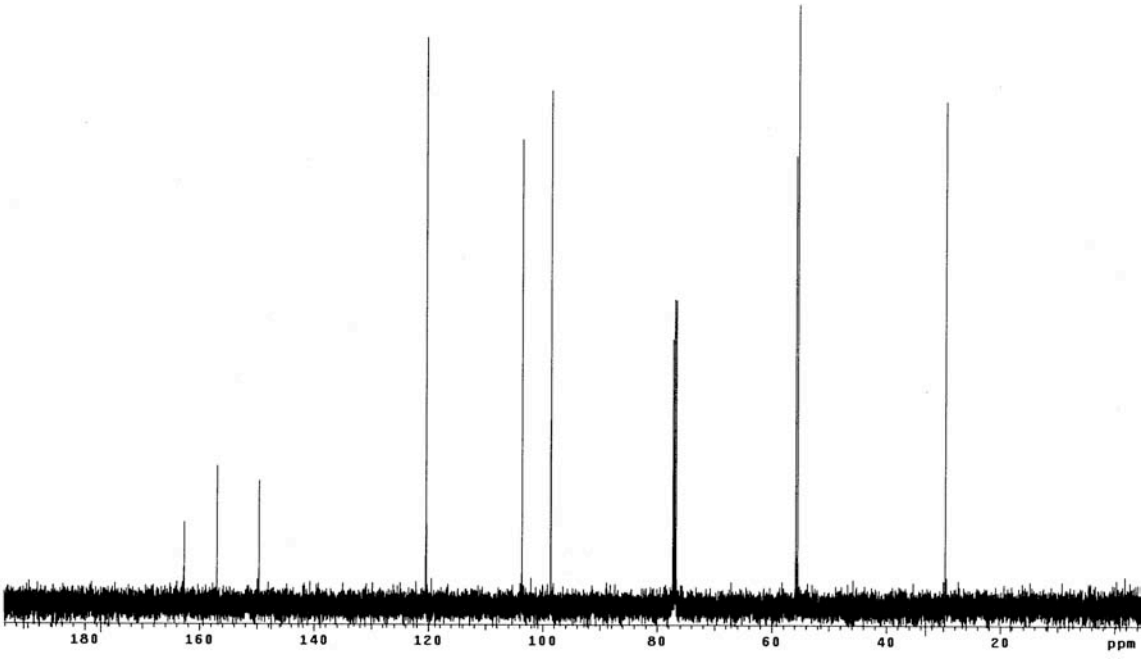
δ = 8.56 (bs, 1H), 8.16 (d, J=8.5, 1H), 6.44-6.48 (m, 2H), 3.99 (s, 2H), 3.86 (s, 3H), 3.78 (s, 3H)

^{13}C NMR (125 MHz, $CDCl_3$)

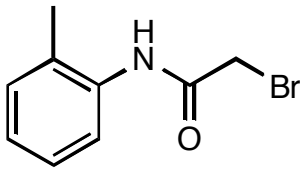
δ = 162.7, 157, 150, 120.4, 103.6, 98.6, 55.8, 55.5, 29.7



^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)



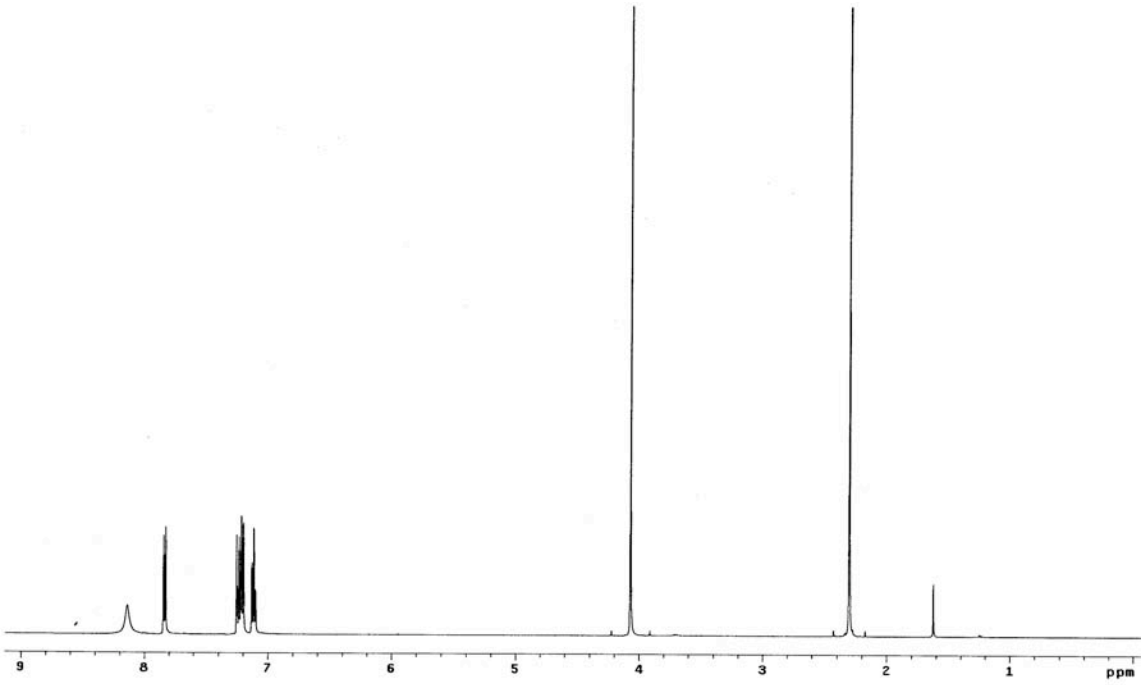
23v

^{1}H NMR (500 MHz, $CDCl_3$)

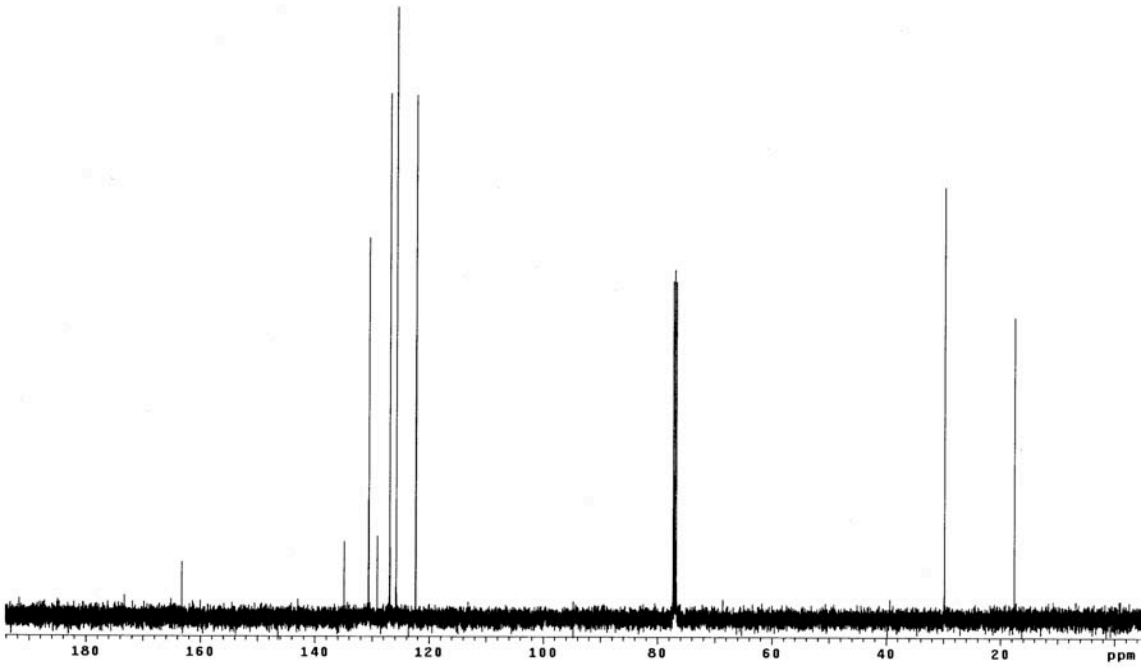
δ = 8.14 (bs, 1H), 7.84 (d, J=8.0, 1H), 7.18-7.26 (m, 2H), 7.12 (t, J=8.0, 1H), 4.07 (s, 2H), 2.3 (s, 3H)

^{13}C NMR (125 MHz, $CDCl_3$)

δ = 163.2, 134.8, 130.6, 129.1, 126.9, 125.8, 122.3, 29.8, 17.6



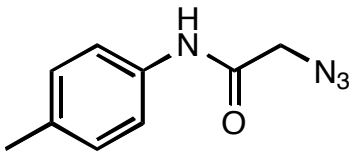
^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

Syntheses of 2-Azido-*N*-Aryl Acetamides 9

Sodium azide (2 equiv) was added to a solution of 2-bromo-*N*-aryl acetamides **23** in acetone/DMF (3:1). The mixture was stirred for 5 h at 40˚C. Complete reaction was observed at this stage by TLC (30% EtOAc in hexane); the acetone was removed, and H_2O (100 mL) was added. The mixture was extracted with ethyl acetate (3x). The organic phase was dried over $MgSO_4$, and was evaporated *in vacuo* to give the corresponding amidoazides **9** in good yields (> 97%). All compounds were verified by their MS, ^1H and ^{13}C NMR. The spectral data of three compounds: **9u**, 9**z** and **9v**, were taken for illustration.



9u

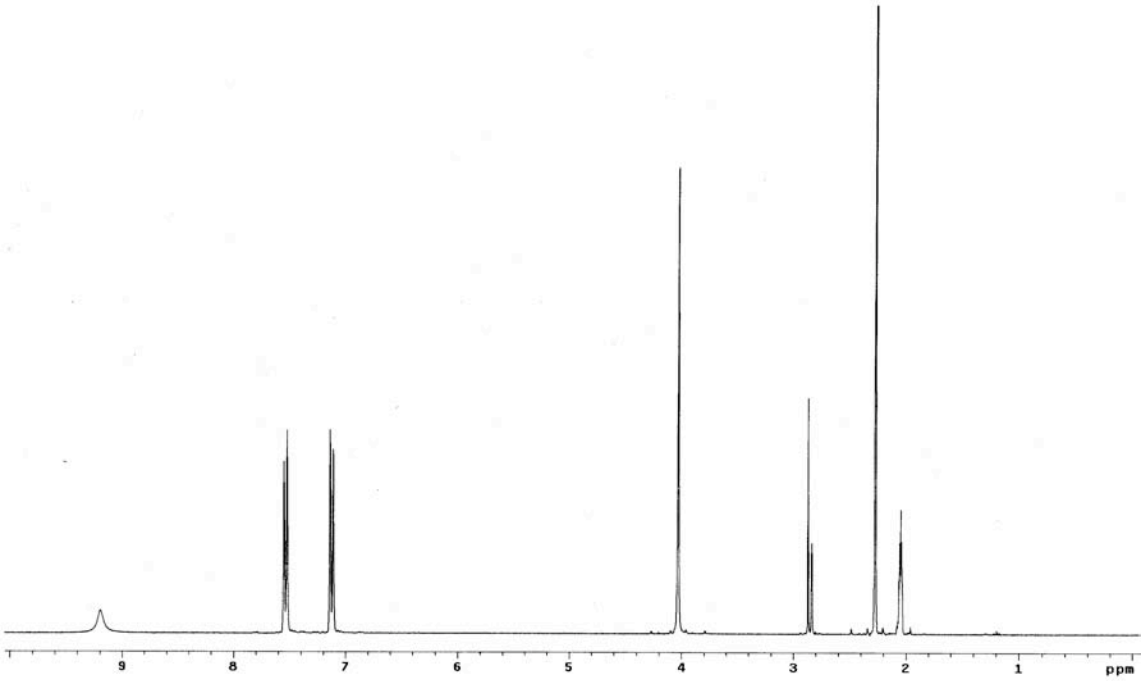
^{1}H NMR (500 MHz, acetone-d6)

δ = 9.20 (bs, 1H), 7.53 (d, J=8.0, 2H), 7.12 (d, J=8.0, 2H), 4.02 (s, 2H), 2.27 (s, 3H)

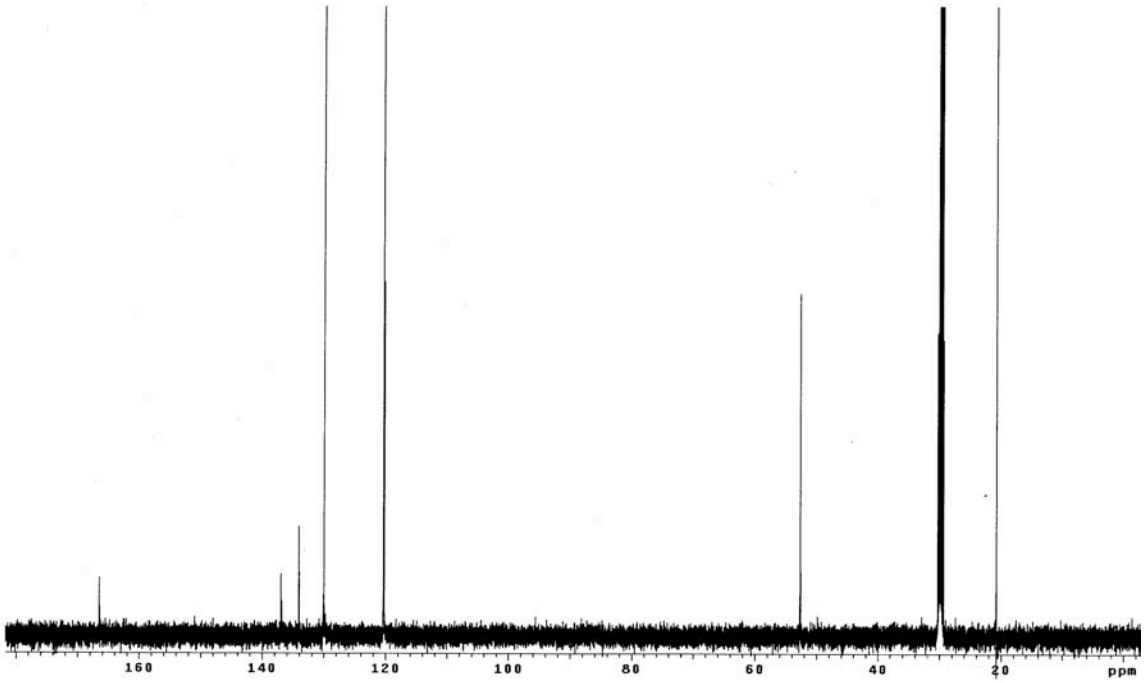
^{13}C NMR (125 MHz, acetone-d6)

δ = 166.5, 136.9, 134.1, 130, 120.3, 52.7, 20.8

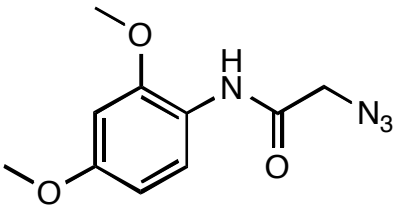
MS (ESI) calc'd for $[C_9H_{10}N_4O + H]^+$ 191.09, found 191.0936



^{1}H NMR (acetone-d6)



^{13}C NMR (acetone-d6)



9z

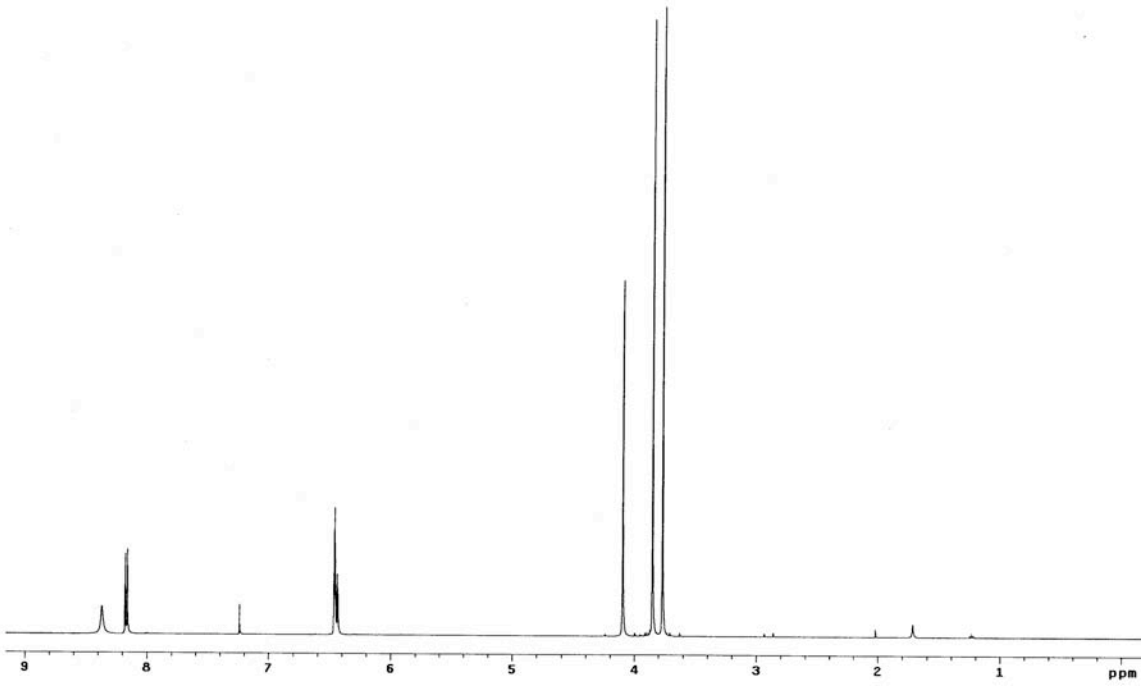
^{1}H NMR (500 MHz, $CDCl_3$)

δ = 8.37 (bs, 1H), 8.17 (d, J=8.5, 1H), 6.45 (m, 2H), 4.10 (s, 2H), 3.85 (s, 3H), 3.77 (s, 3H)

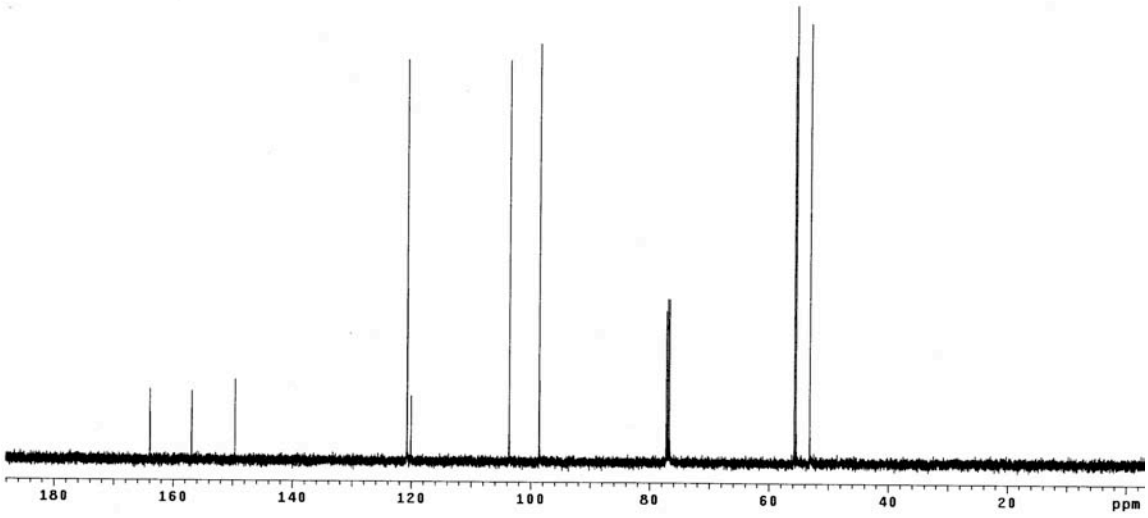
^{13}C NMR (125 MHz, $CDCl_3$)

δ = 163.9, 156.9, 149.6, 120.7, 120.1, 103.6, 98.6, 55.8, 55.5, 53.2

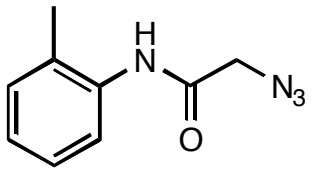
MS (ESI) calc'd for $[C_{10}H_{12}N_4O_3 + H]^+$ 237.10, found 237.1068



^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)



9v

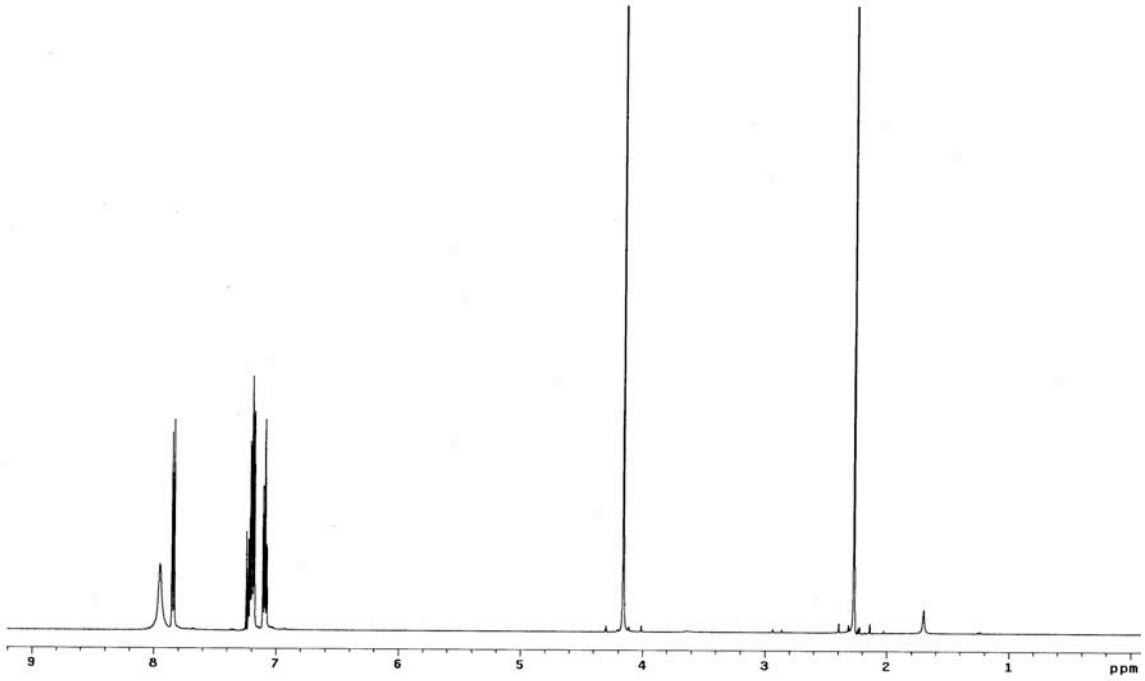
^{1}H NMR (500 MHz, $CDCl_3$)

 = 7.95 (bs, 1H), 7.84 (d, J=8.5, 1H), 7.17-7.23 (m, 2H), 7.09 (t, J=7.5, 1H), 4.16 (s, 2H), 2.27 (s, 3H)

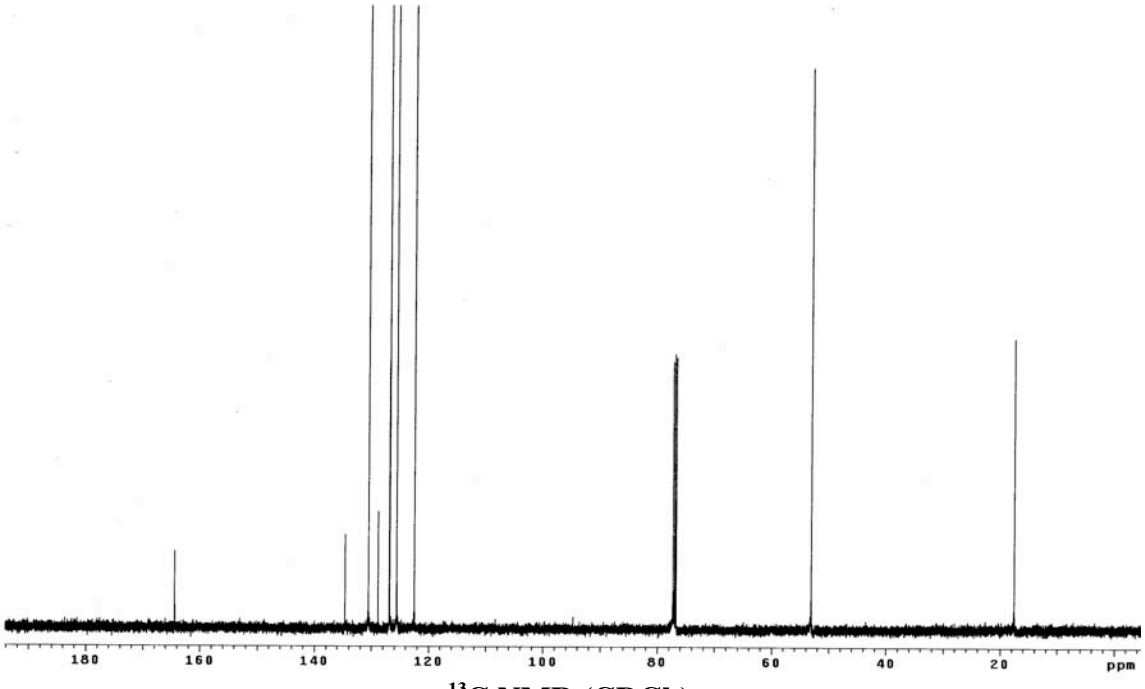
^{13}C NMR (125 MHz, $CDCl_3$)

 = 164.4, 134.6, 130.5, 128.8, 126.8, 125.6, 122,5, 53.1, 17.5

MS (ESI) calc'd for $[C_9H_{10}N_4O + H]^+$ 191.09, found 191.1004



^{1}H NMR ($CDCl_3$)



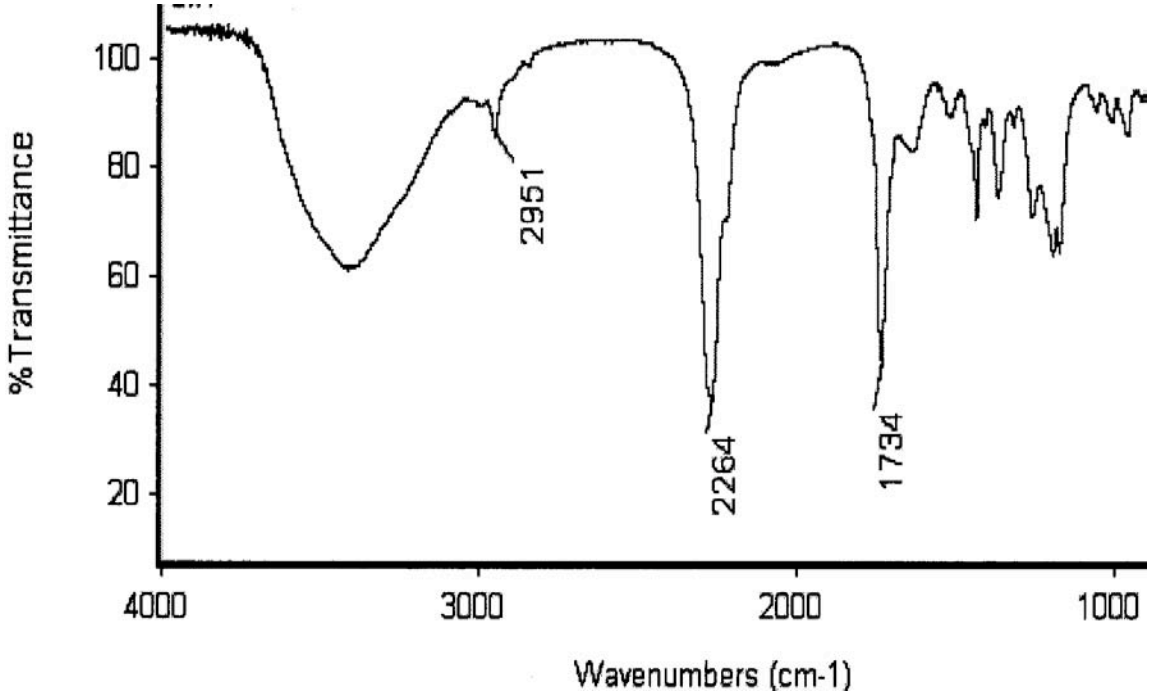
^{13}C NMR ($CDCl_3$)

Syntheses of Amino Acid Ester Isocyanates 24 and Propargyl Isocyanate 8a

A saturated 100ml aqueous solution of $NaHCO_3$ was added to a solution of □-alanine methyl ester hydrochloride or L-phenylalanine methyl ester hydrochloride or propargylamine (21.8 mmol) in 100 mL CH_2Cl_2. The biphasic mixture was cooled in an ice bath and stirred mechanically, then triphosgene (2.16 g, 7.3 mmol) was added in a single portion. The reaction mixture was stirred vigorously for 20 min at 0 ˚C. The organic layer was separated, and the aqueous layer was extracted with 3 x 20 mL CH_2Cl_2. The combined organic layers were dried over Na_2SO_4, filtered and concentrated to 50 ml for **24** *in vacuo* and 3 ~ 4 ml for **8a** by distillation under atmospheric pressure. The products were characterized by IR.

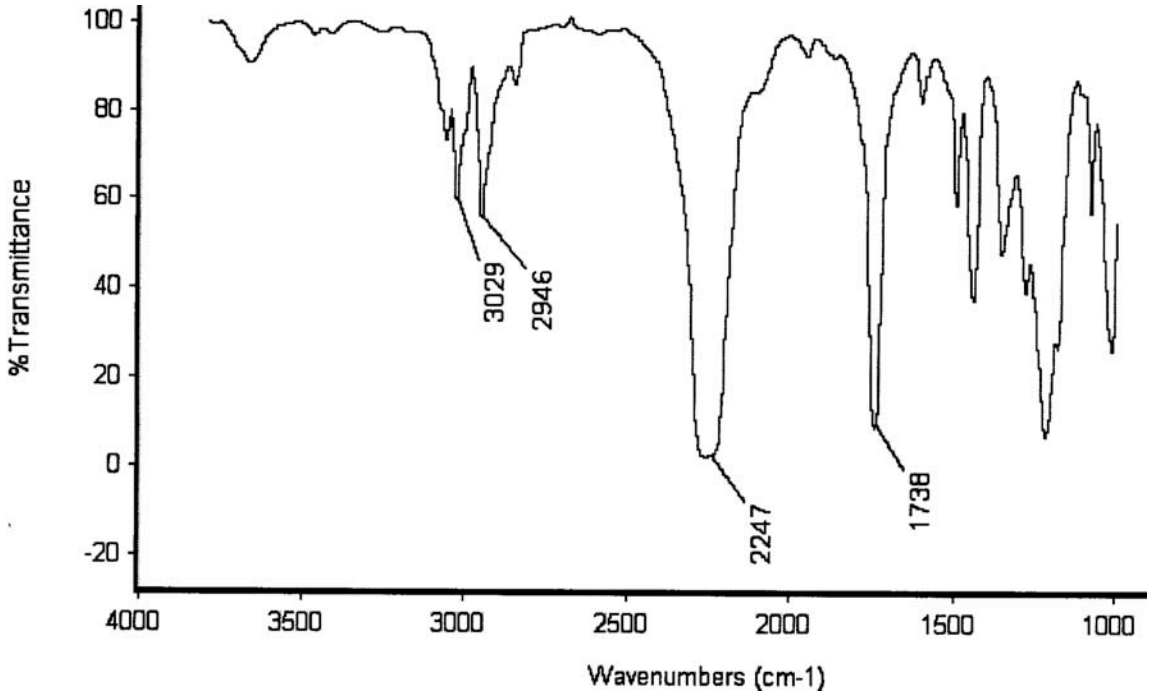
3-isocyanato-propionic acid methyl ester (**24a**):

IR: 2264 cm^{-1} (NCO)



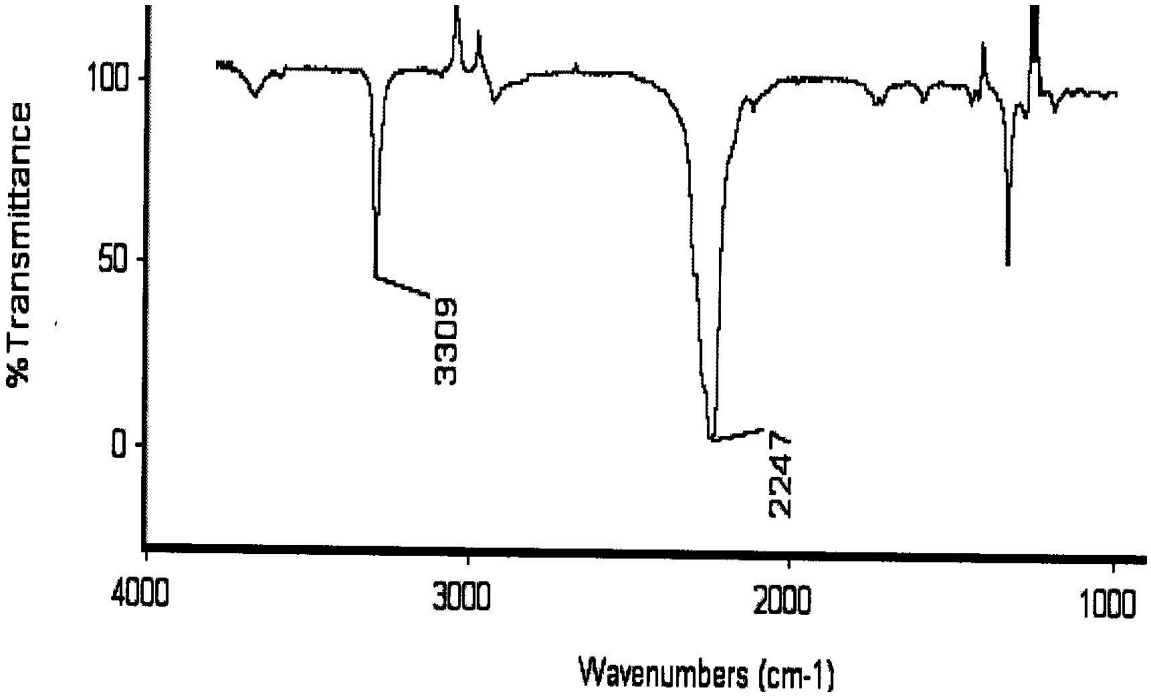
(s)-2-isocyanato-3-phenyl-propionic acid methyl ester (**24b):**

IR: 2247 cm^{-1} (NCO)



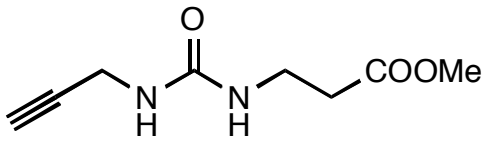
propargyl isocyanate **(8a):**

IR: 2247 cm^{-1} (NCO)



Syntheses of Intermediates 25 and 26

Propargylamine (0.9 equiv) was added to the solution of amino acid ester isocyanates **24** in CH_2Cl_2. The reaction mixture was stirred for 30 min at 25 ˚C. The solvent was removed *in vacuo*. The crude products were chromatographed using EtOAc/Hexane (1:1) eluant to give intermediates **25** in 94% yield 94 % and **26** in 95% yield for two steps. Both are white powder.



25

Mp: 83-84 ˚C

R_f=0.38 (EtOAc/Hexane = 3:1)

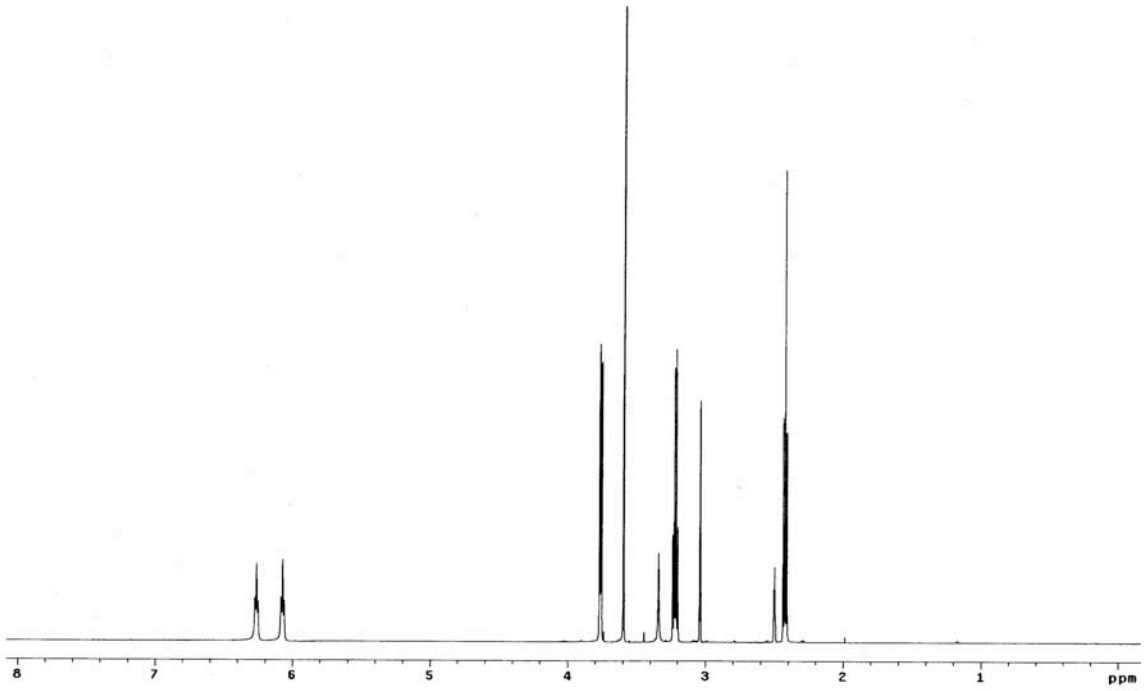
^{1}H NMR (500 MHz, DMSO-d6)

 = 6.26 (t, J=5.5, 1H), 6.07 (t, J=5.5, 1H), 3.76 (dd, J=2.55, 5.50, 2H), 3.59 (s, 3H), 3.22 (q, J=5.5, 2H), 3.04 (t, J=2.0, 1H), 2.42 (t, J=7.0, 2H)

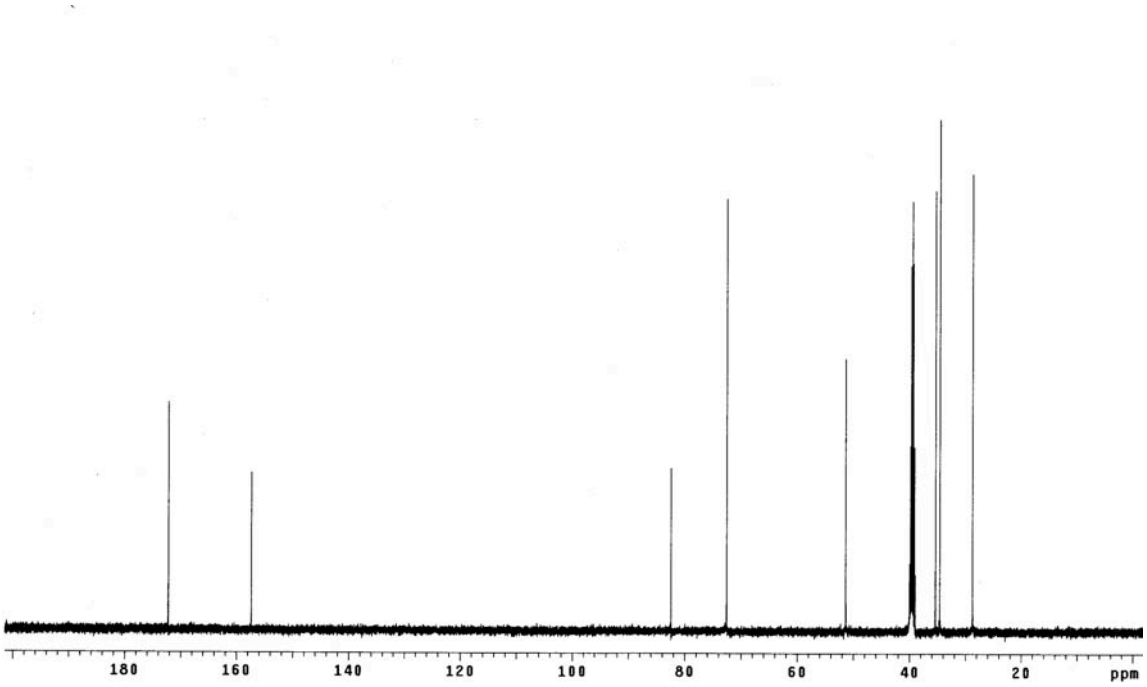
^{13}C NMR (125 MHz, DMSO-d6)

 = 172.2, 157.3, 82.5, 72.5, 51.3, 35.4, 34.6, 28.7

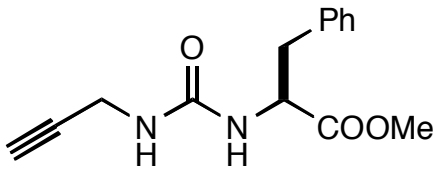
MS (ESI) calc'd for $[C_8H_{12}N_2O_3+H]^+$ 185.08, found 185.09



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



26

Mp: 96 ˚C

R_f=0.43 (EtOAc/Hexane = 1:1)

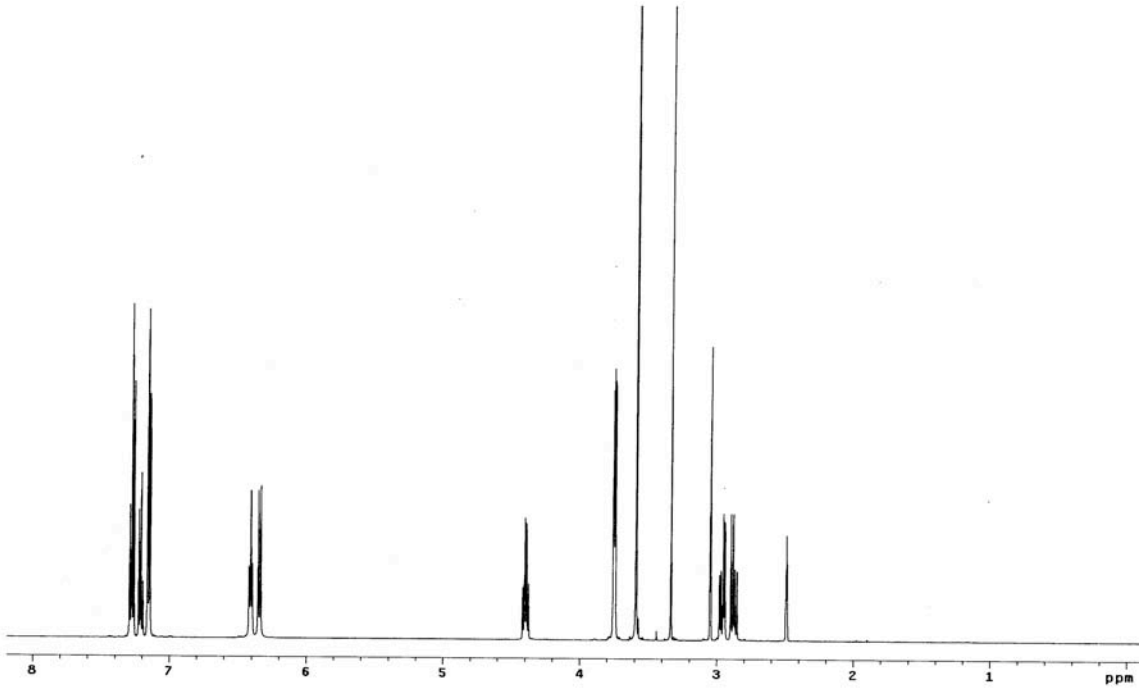
^{1}H NMR (500 MHz, DMSO-d6)

= 7.26-7.30 (m, 2H), 7.19-7.23 (m, 1H), 7.15 (d, J=8.0, 2H), 6.41 (t, J=6.0, 1H), 6.34 (d, J=8.5, 1H), 4.37-4.42 (m, 1H), 3.75 (dd, J=2.45, 6.0, 2H), 3.59 (s, 3H), 3.05 (t, J=2.45, 1H), 2.96 (dd, J=5.5, 14.0, 1H), 2.88 (dd, 5.5, 14.0, 1H)

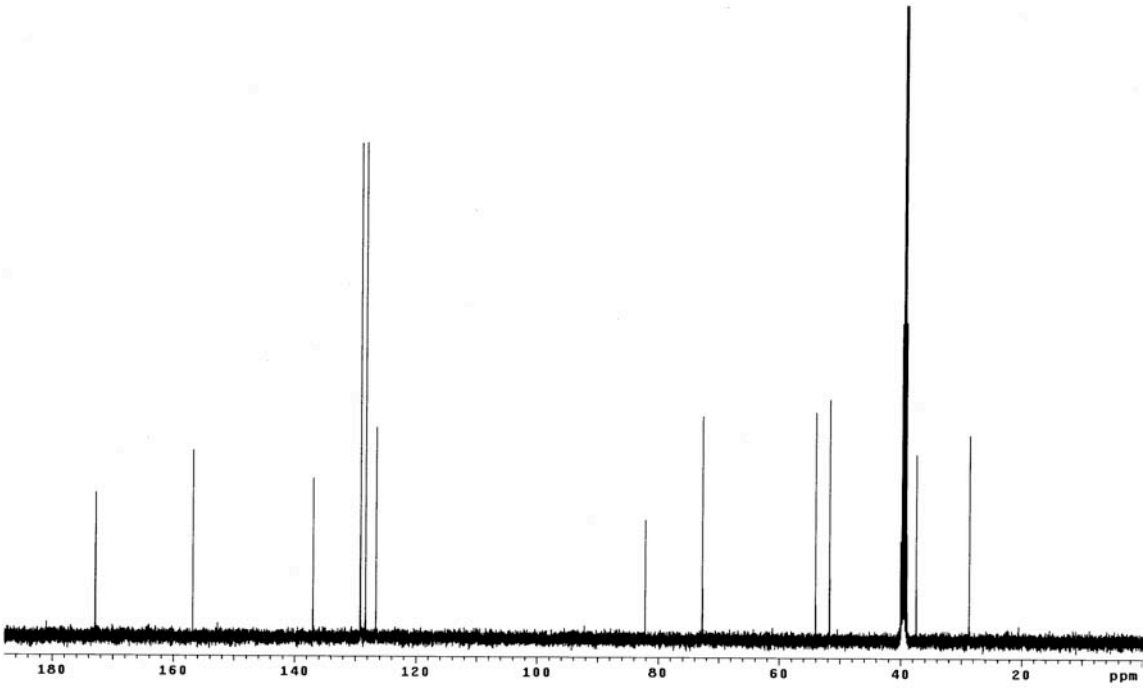
^{13}C NMR (125 MHz, DMSO-d6)

= 172.9, 156.8, 137, 129.2, 128.3, 126.6, 82.2, 72.7, 54, 51.7, 37.5, 28.7

MS (ESI) calc'd for $[C_{14}H_{16}N_2O_3+Na]^+$ 283.11, found 283.1096



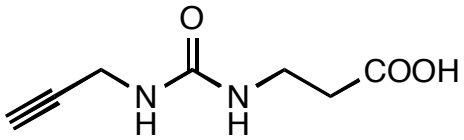
^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

Syntheses of Alkyne Building Blocks 8b and 8c

Lithium hydroxide monohydrate (1.3 equiv) was added to a solution of compound **25** or **26** (1 equiv) in THF/MeOH/H_2O (3:2:1) (about 1.7mL the mixture solvents per mmol **25** or **26**). The mixture was stirred at 25˚C for 2 h. The solvent was removed *in vacuo*, the residue was diluted with water and then acidified with 2M HCl up to a pH of 2-3. The resulting solution was extracted with 6 x 30 mL ethyl acetate, dried over anhydrous sodium sulfate. The dried solution was concentrated *in vacuo* to give products **8b-c** in 99% yield.



8b

Mp: 132-133 ˚C

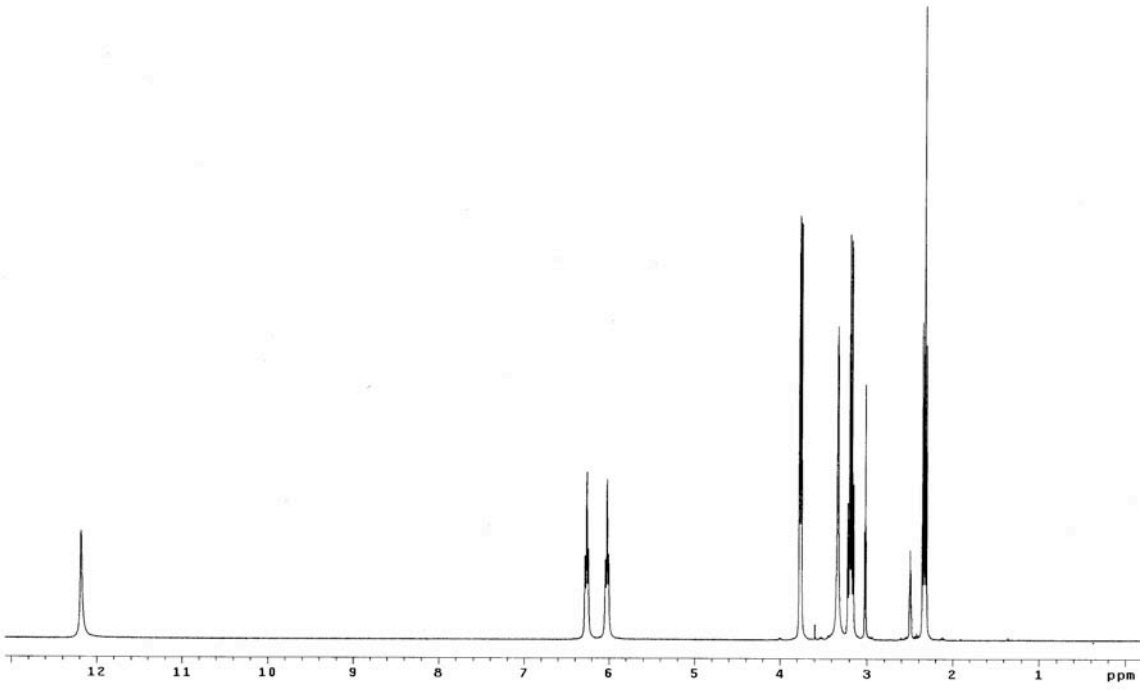
^{1}H NMR (300 MHz, DMSO-d6)

δ = 12.18 (s, 1H), 6.26 (t, J=5.4, 1H), 6.03 (t, J=5.4, 1H), 3.76 (dd, J=2.46, 5.70, 2H), 3.19 (q, J=6.0, 2H), 3.04 (t, J=2.55, 1H), 2.33 (t, J=6.6, 2H)

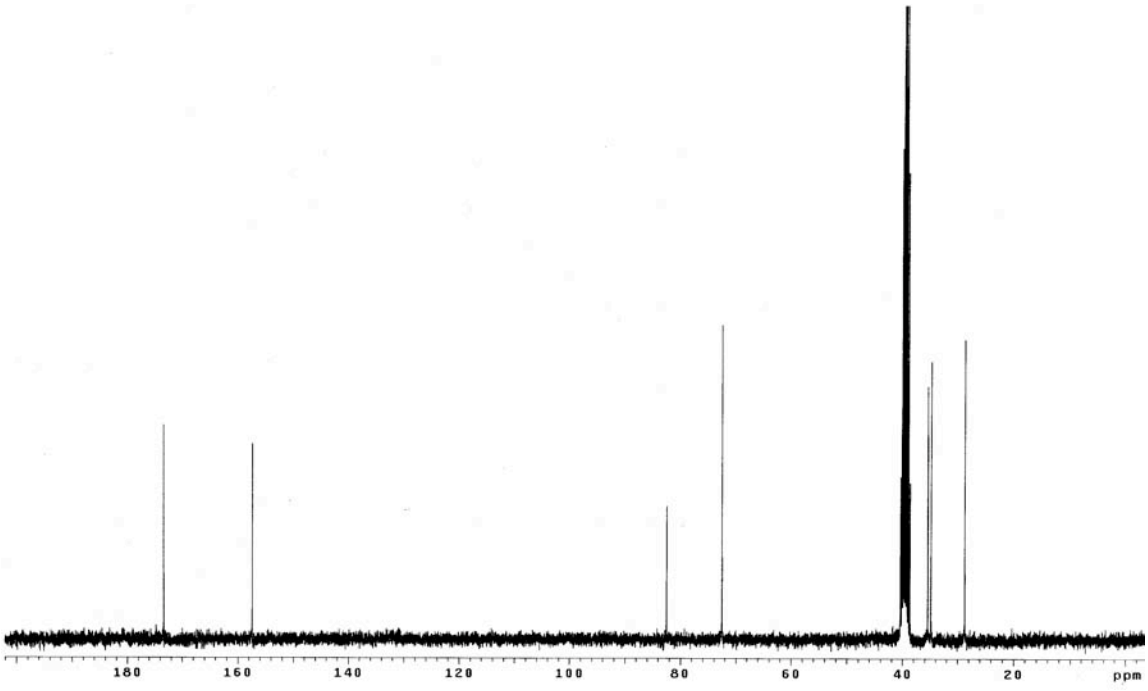
^{13}C NMR (75 MHz, DMSO-d6)

δ = 173.3, 157.3, 82.5, 72.5, 35.4, 34.8, 28.7

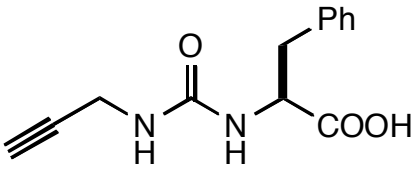
MS (ESI) calc'd for $[C_7H_{10}N_2O_3+H]^+$ 171.07, found 171.0795



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



8c

Mp: 152 ˚C

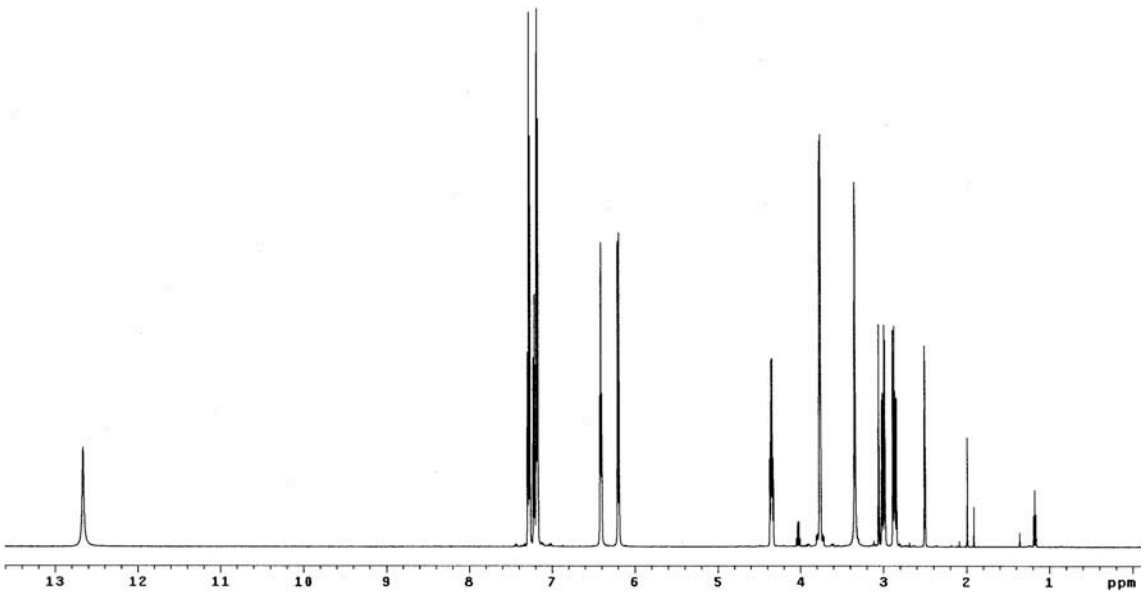
^{1}H NMR (500 MHz, DMSO-d6)

 = 12.67 (s, 1H), 7.25-7.29 (m, 2H), 7.15-7.22 (m, 3H), 6.40 (t, J=5.5, 1H), 6.19 (d, J=8.0, 1H), 4.31-4.37 (m, 1H), 3.74-3.77 (m, 2H), 3.59 (s, 3H), 3.04 (t, J=2.45, 1H), 2.99 (dd, J=5.5, 14.0, 1H), 2.86 (dd, 5.5, 14.0, 1H)

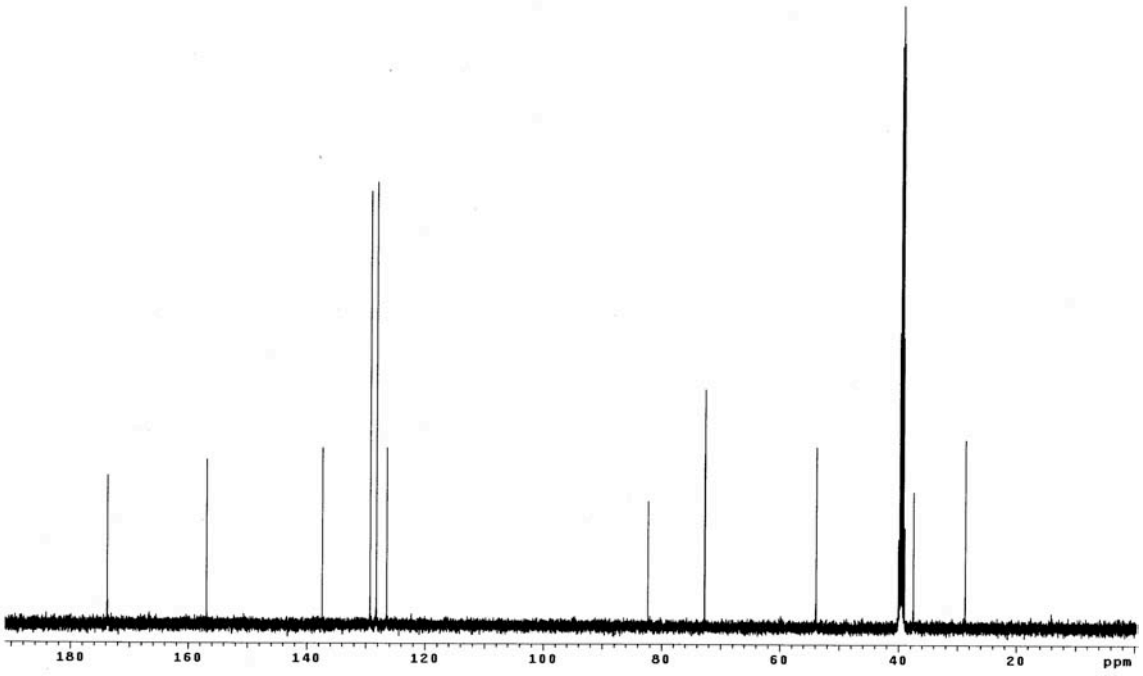
^{13}C NMR (125 MHz, DMSO-d6)

 = 173.8, 156.9, 137.4, 129.3, 128.2, 126.5, 82.3, 72.7, 53.9, 37.5, 28.7

MS (ESI) calc'd for $[C_{13}H_{14}N_2O_3\text{-}H]^+$ 245.09, found 245.0772



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

Preparation of Solid Phase Supported Poly(propylene imine) Dendrimers 28 and 30

TentaGel S PHB with loading 0.24 mmol/g was swelled with CH_2Cl_2 in a fritted syringe for 30 min. 4-nitrophenylchloroformate (6 equiv) and N-methylmorpholine (8 equiv) in CH_2Cl_2 was added. After gentle shaking for 24 h, the reaction mixture was then drained, and the resin was washed with DMF (3x), MeOH (3x), CH_2Cl_2 (3x), MeOH (3x), CH_2Cl_2 (3x). The resin containing the linker was treated with DAB-Am-4 (or 8) (5 equiv), 4-(dimethylamino)pyridine (3 equiv) in DMF/CH_2Cl_2 2:1 (or 1:1) for 48 h. The mixture was washed with Et_3N (3x), MeOH (3x), CH_2Cl_2 (3x), DMF (3x), MeOH (3x), CH_2Cl_2 (3x) and dried *in vacuo* for later use.

General Procedure for Preparation of Functionalized Poly(propylene imine) Dendrimers 29 and 31

TentaGel supported poly(propylene imine) dendrimer **28** or **30** was swelled with CH_2Cl_2 for 30 min. The resin-supported DAB dendrimer was treated with functional

compounds **15**-**21**, HOBt, HBTU, and DIEA in DMF/CH_2Cl_2 (1.5:1 for **29**, 1:1 for **31**). After gentle shaking for 24 h, the reaction mixture was drained and the resin was washed with DMF (3x), MeOH (3x), CH_2Cl_2 (3x), DMF (3x), MeOH (3x), CH_2Cl_2 (3x). The functionalized DAB dendrimer was cleaved from the resin by treating with a mixture of 50% TFA, 5% TIS, and 45% CH_2Cl_2 for 3 h. The cleavage solution was separated from the resin via filtration. After most of the cleavage cocktail was evaporated in a stream of nitrogen, the crude functionalized DAB dendrimer was triturated using cooled anhydrous ethyl ether. The purity of this crude product was determined by analytical HPLC (SSI system, 5-95% B in 30 min). The crude product was dissolved in the mixture of CH_3CN and H_2O, filtered, then purified with preparative HPLC and finally lyophilized to yield the final product as white powder.

29a

^{1}H NMR (300 MHz, CD_3CN)

δ = 7.83 (m, 6H), 7.74 (m, 3H), 7.39 (m, 6H), 7.21 (m, 11H), 6.61 (m, 3H), 4.49 (m, 3H), 3.26 (m, 1H), 3.17 (m, 2H), 3.03-3.06 (m, 5H), 2.98 (m, 4H), 2.93 (m, 2H), 2.87 (m, 3H), 2.81 (m, 6H), 2.37 (s, 9H), 1.97 (m, 6H), 1.85 (m, 6H), 1.74 (m, 4H), 1.66 (m, 10H), 1.57 (m, 4H)

^{13}C NMR (75 MHz, CD_3CN)

δ = 175.4, 163.4, 147.6, 135.8, 131.4, 128.1, 125.6, 123.1, 122.4, 113.2, 70.7, 53.8, 52.5, 51.3, 50.8, 50.2, 49.5, 45.5, 37.8, 36.6, 33.1, 28.5, 24.5, 24.4, 23.4, 22.4, 21.7, 21.3

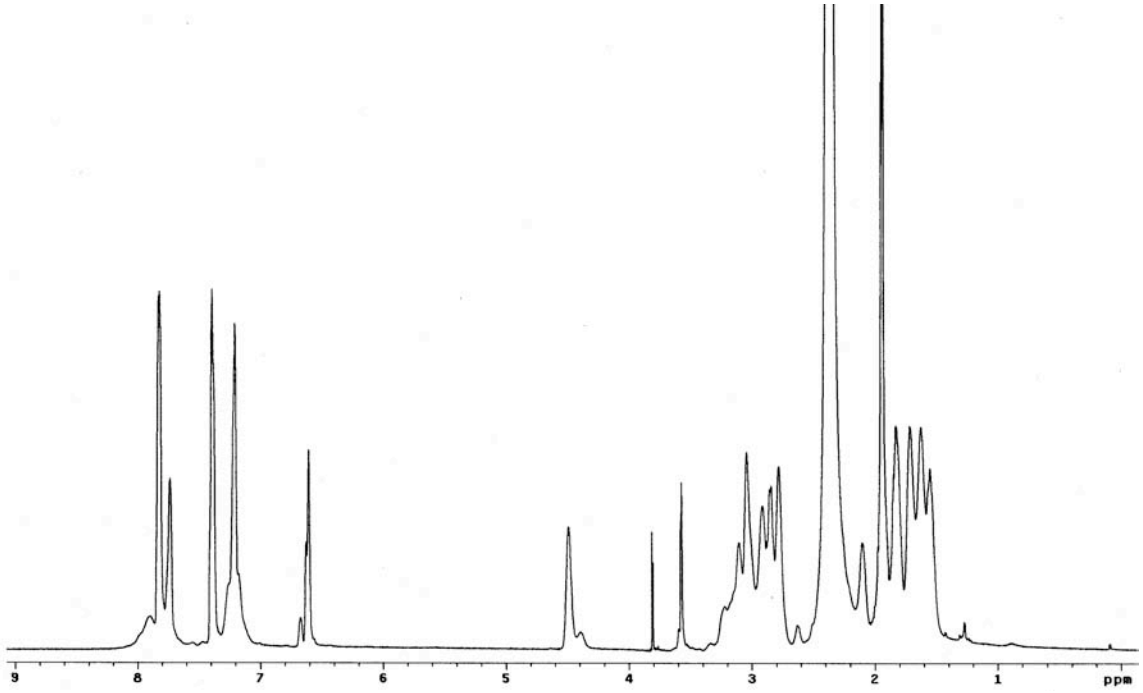
IR (CH_3CN as solvent)

3274-3596 (N-H and NH_2), 1635 (CO)

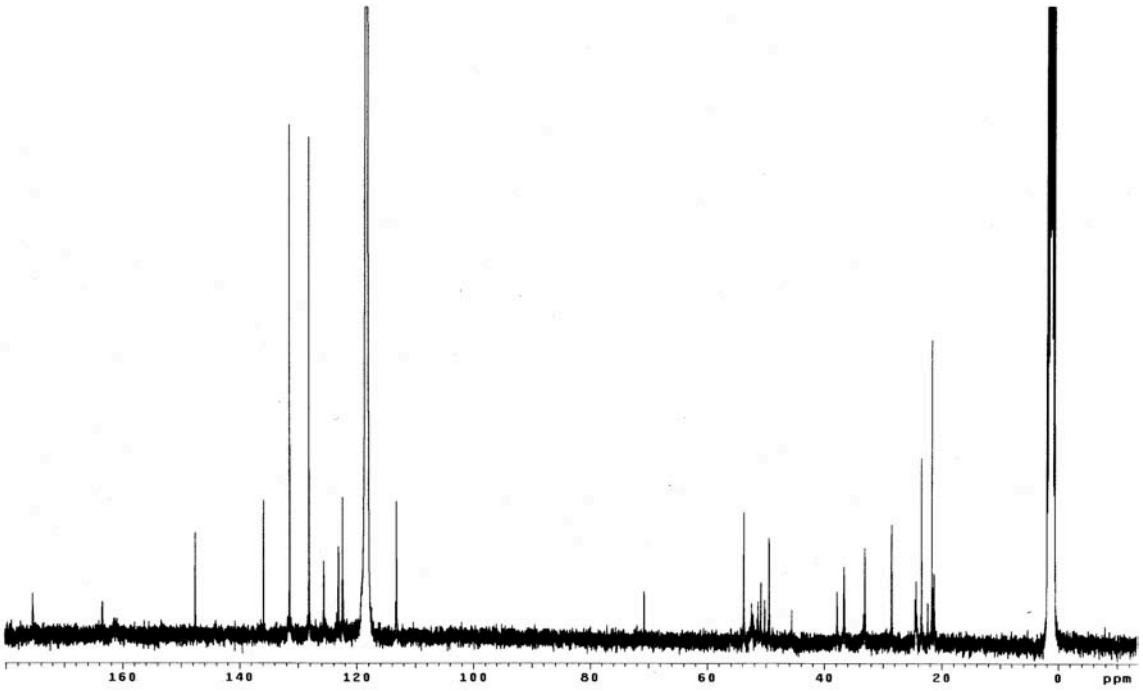
MS (MALDI) calc'd for $[C_{70}H_{94}N_{12}O_{12}S_3Na]^+$ 1414.75, found 1414.66

Analytical HPLC

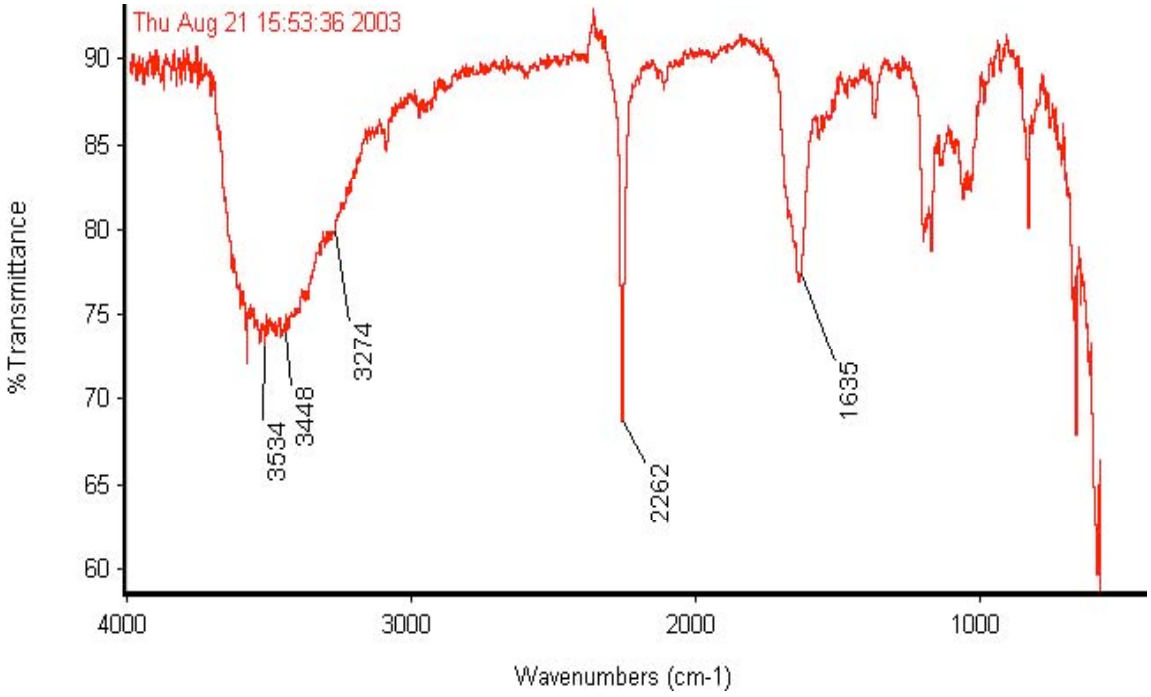
homogeneous single peak, retention time = 19.7 min (5-95% B in 35 min)



^{1}H NMR (CD_3CN)



^{13}C NMR (CD_3CN)



IR (CH_3CN as solvent)

29b

^{1}H NMR (300 MHz, CD_3CN)

δ = 7.89 (m, 6H), 7.74 (m, 3H), 7.21 (m, 11H), 7.05 (m, 6H), 6.62 (m, 3H), 4.50 (m, 3H), 3.84 (s, 9H), 3.26 (m, 1H), 3.17 (m, 2H), 3.03-3.06 (m, 5H), 2.97 (m, 4H), 2.93 (m, 2H), 2.87 (m, 3H), 2.80 (m, 6H), 1.96 (m, 6H), 1.85 (m, 6H), 1.74 (m, 4H), 1.65 (m, 10H), 1.57 (m, 4H)

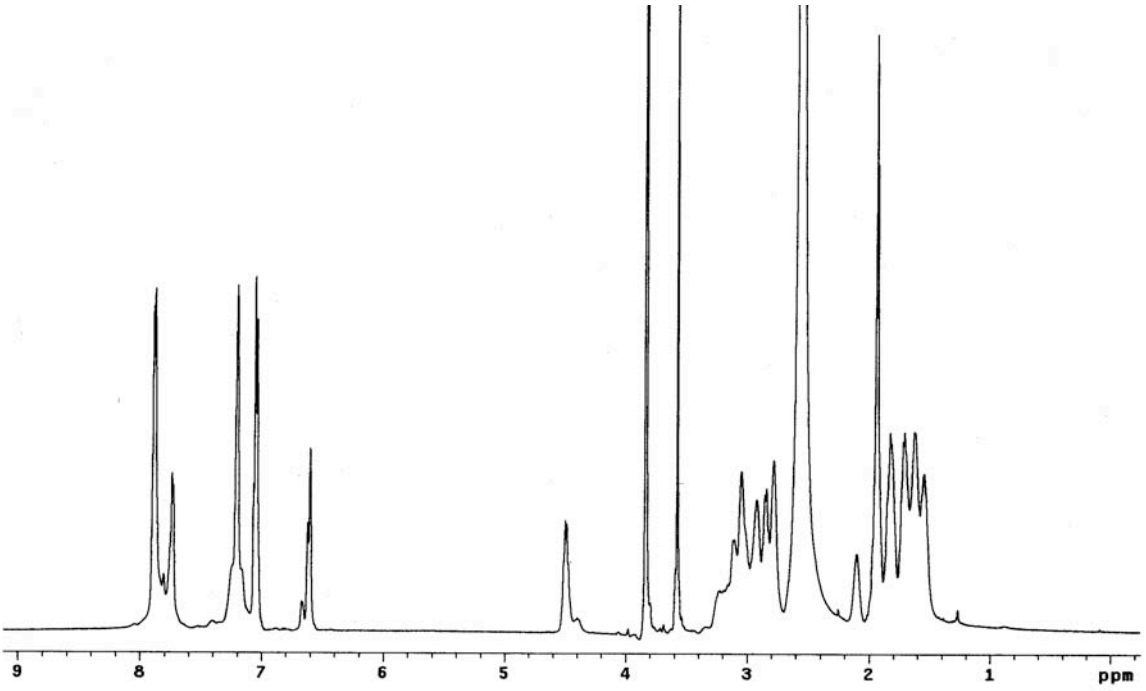
^{13}C NMR (75 MHz, CD_3CN)

δ = 175.5, 165.7, 163.6, 130.7, 130.0, 125.5, 123.2, 122.4, 116.2, 113.2, 70.9, 56.9, 53.9, 52.6, 51.0, 50.4, 49.6, 37.9, 36.7, 33.2, 28.6, 24.6, 23.5, 22.5, 21.4

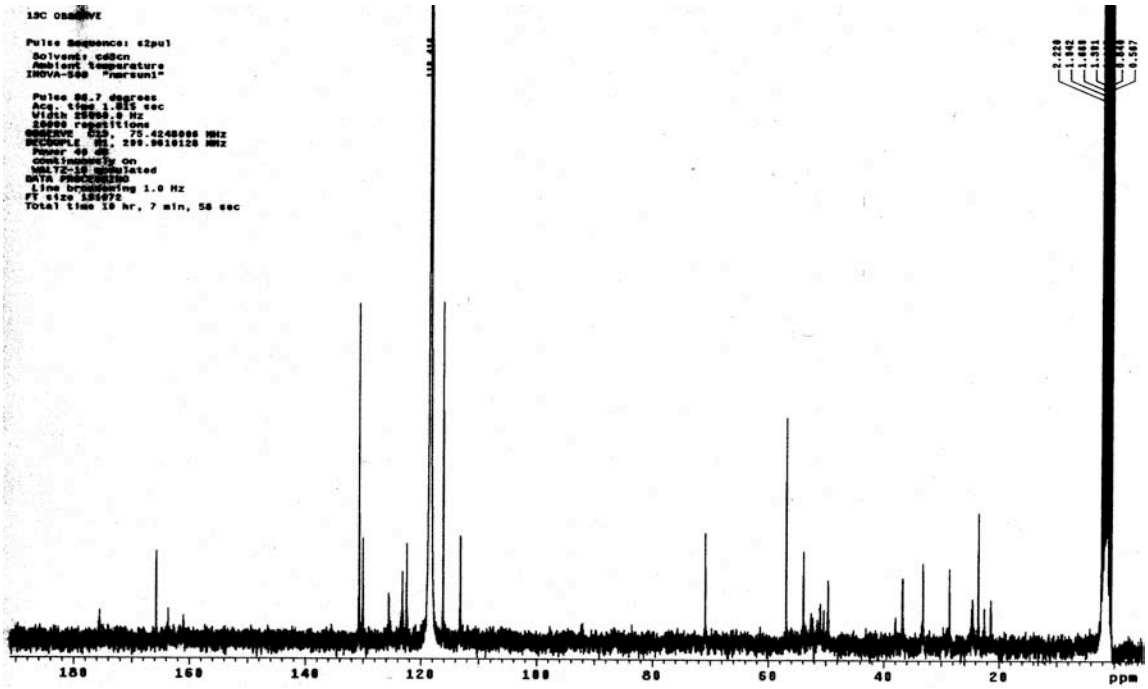
MS (MALDI) calc'd for $[C_{70}H_{94}N_{12}O_{15}S_3Na]^+$ 1462.75, found 1462.45

Analytical HPLC

homogeneous single peak, retention time = 18.2 min (5-95% B in 35 min)



^{1}H NMR (CD_3CN)



^{13}C NMR (CD_3CN)

31

^{1}H NMR (300 MHz, DMSO-d6)

δ = 7.92 (s, 7H), 7.84 (m, 30H), 7.44 (m, 14H), 7.32 (m, 7H), 6.66 (m, 7H), 4.40 (m, 7H), 2.71-3.16 (m, 59H), 2.35 (s, 21H), 1.51-2.05 (m, 63H), 1.44 (m, 7H)

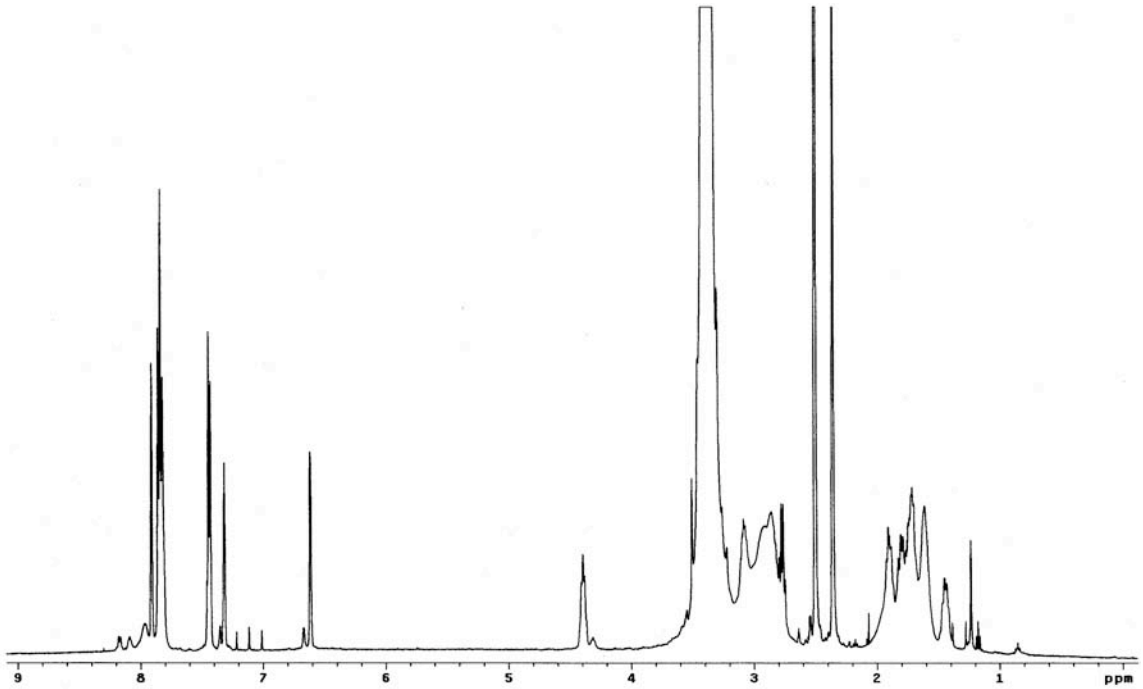
^{13}C NMR (75 MHz, DMSO-d_6)

δ = 172.8, 161.6, 146.0, 134.5, 130.4, 126.9, 124.5, 121.9, 121.1, 112.9, 69.8, 52.3, 50.0, 47.7, 35.7, 31.4, 27.4, 23.3, 22.5, 21.1

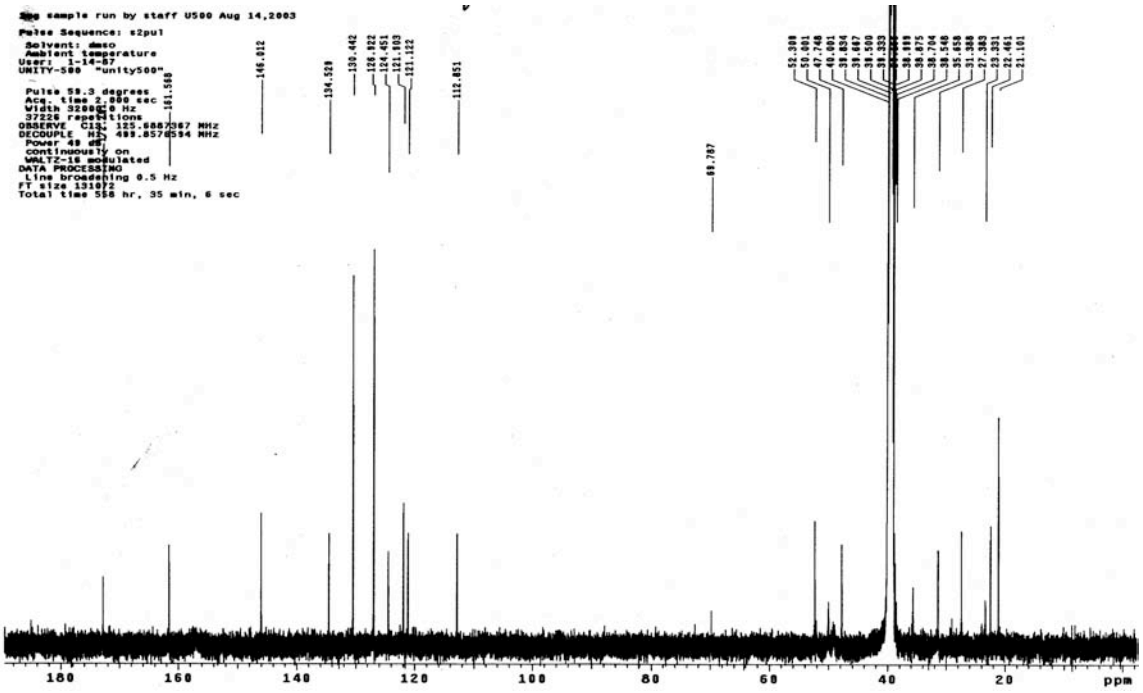
MS (MALDI) calc'd for $[C_{166}H_{222}N_{28}O_{28}S_7Na]^+$ 3302.48, found 3302.46

Analytical HPLC

homogeneous single peak, retention time = 23.5 min (5-95% B in 35 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

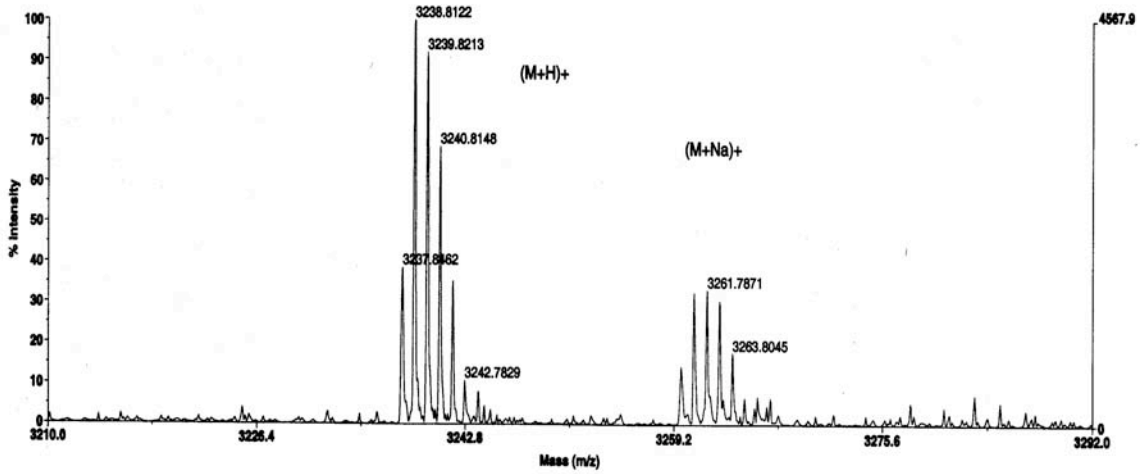
Procedure for Preparation of Amino Acid Modified Poly(propylene imine) Dendrimer 32

TentaGel supported poly(propylene imine) dendrimer **28** was swelled with CH_2Cl_2 for 30 min. The following cycle was then followed: the resin-supported DAB dendrimer was treated with Fmoc-His(Trt)-OH (4 equiv), HOBt (4 equiv), DIC (4 equiv), and DIEA (6 equiv) in DMF/CH_2Cl_2 (1:1) for 24 h, the Fmoc protecting group was removed by treating the resin twice with 20% piperidine in DMF (10 min and 15 min). The resin was washed with DMF (3x), CH_3OH (3x) and CH_2Cl_2 (3x). After washing, the resin was treated with Fmoc-Leu-OH and coupling reagents for 24 h. The above cycle was repeated seven times, each of which the desired Fmoc protected amino acid was used until Asp was coupled. Finally succinic anhydride was used to cap N-terminus, the work-up and purification are the same as those for **29** and **31**. The amino acid sequence is the same as that of peptidomimetic **6** discussed in Chapter III.

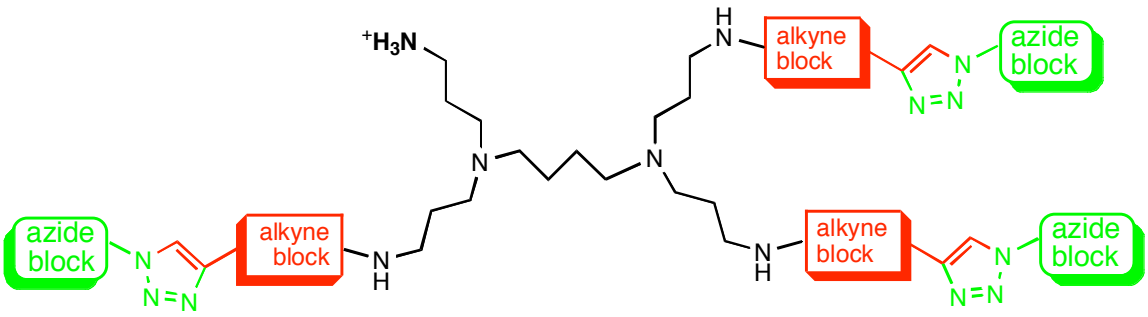
MS (MALDI) calc'd for $[C_{160}H_{229}N_{33}O_{39}+H]^+$ 3239.8, found 3239.8

Analytical HPLC

homogeneous single peak, retention time = 16.35 min (5-95% B in 35 min)



MS (MALDI)

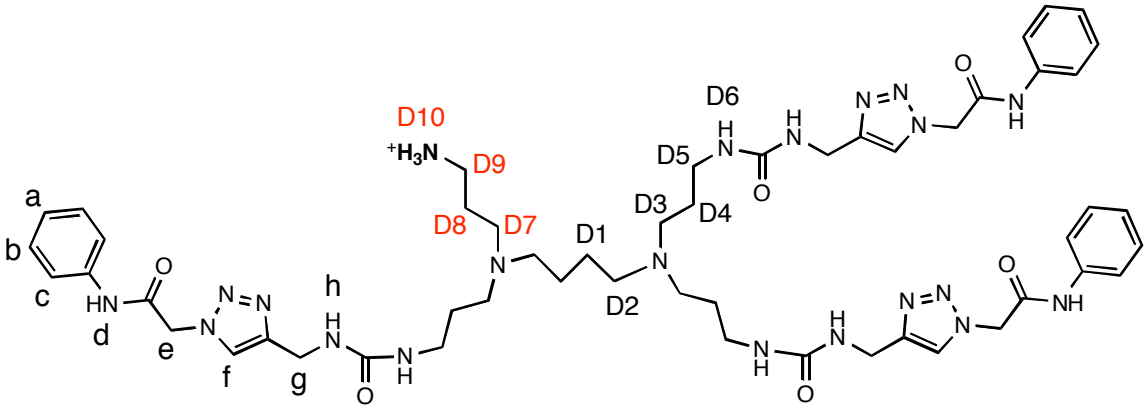


10

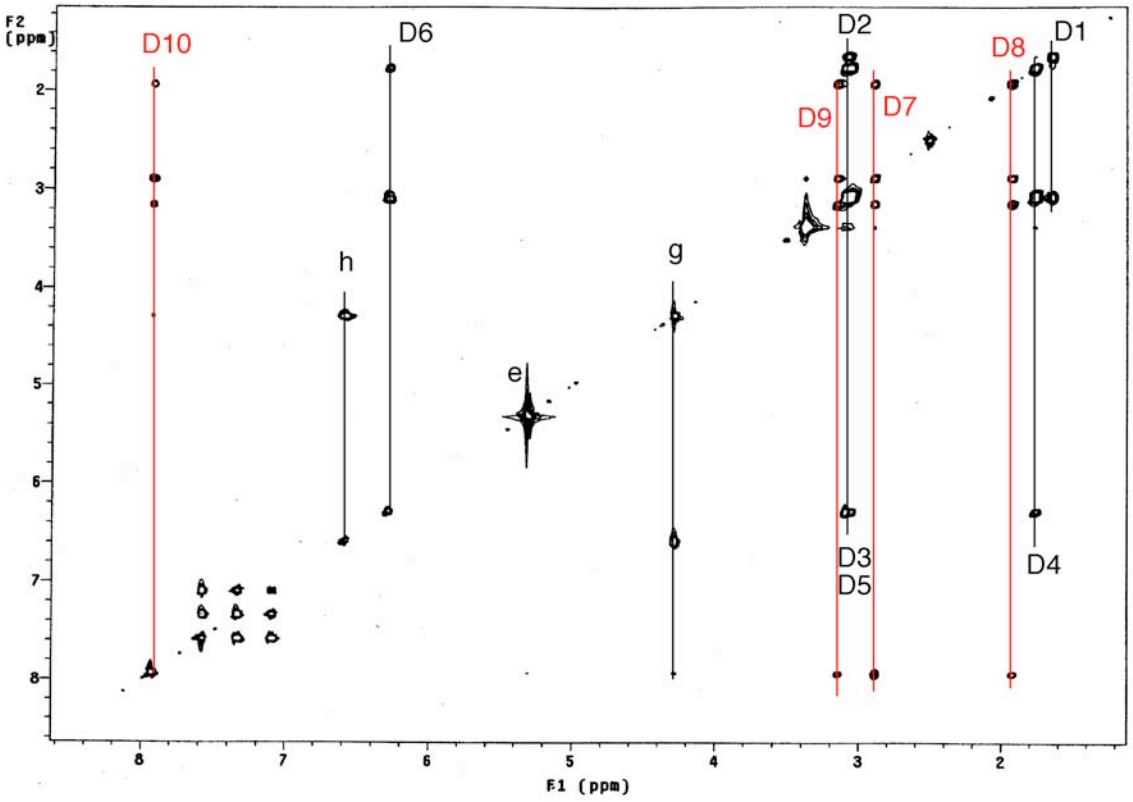
General Procedure for Preparation of Trivalent Protein G Mimics 10

The resin with activated carbonate **27** was swelled with CH_2Cl_2 in a fritted syringe for 30 min. Poly(propylene imine) dendrimer (DAB-Am-4) was then coupled to activated carbonate **27** using the following conditions: 1.5 M DAB-Am-4 and a catalytic amount of 4-(dimethylamino)pyridine in DMF/CH_2Cl_2 (2:1) at 25 ˚C for 48 h. The TentaGel Supported dendrimer **28** was treated with alkyne building block **8a-c**, HOBt, HBTU, and DIEA in DMF/CH_2Cl_2 (1.5:1), After gentle shaking for 18 h at 25˚C, the reaction mixture was drained, and the resin was washed via two cycles of DMF (3x), MeOH (3x), CH_2CL_2 (3x). The resin was dried over night, and divided into many portions as required (*e.g.* 4 portions for **8a**, 6 portions for **8b-c**). Each solid supported alkyne portion was swelled again, and then treated with azide building blocks **9**, CuI (2.5 equiv) and DIEA (40 equiv) in THF at 25 ˚C for 20 h. The reaction mixture was drained and the resin was washed with THF (3x), H_2O (3x), then DMF (2x), MeOH (2x), CH_2Cl_2 (3x). The protein G mimic was cleaved from the resin by treating with a mixture of 50% TFA in CH_2Cl_2 for 2 h. The cleavage solution was separated from the resin via filtration. After most of the cleavage cocktail was evaporated in a stream of nitrogen, the crude product was triturated using cooled anhydrous ethyl ether. All crude materials were analyzed by analytical HPLC (SSI system, 5-95% B in 30 min) and MALDI-MS. The selected crude materials were purified via preparative HPLC and lyophilized to yield the

desired products. Spectral data were obtained: ^{1}H NMR for all compounds, ^{13}C NMR for eight compounds, TOCSY (mix=80ms) for three cmpounds, which have different alkyne building blocks.



10at



80-ms TOCSY of 10at (DMSO-d6)

^{1}H NMR (500 MHz, DMSO-d6)

δ = 10.5 (s, 3H, C_6H_5NHCO), 7.93 (bs, 2H, $NCH_2CH_2CH_2NH_2$), 7.92 (s, 3H), 7.57 (d, J=7.5, 6H), 7.33 (t, J=7.5, 6H), 7.09 (t, J=7.5, 3H), 6.55-6.63 (m, 3H, $NHCONHCH_2CH_2CH_2N$), 6.25-6.32 (m, 3H, $NHCONHCH_2CH_2CH_2N$), 5.3 (s, 6H), 4.26-4.30 (m, 6H), 3.0-3.18 (m, 18H, $NCH_2CH_2CH_2NH_2$, $NCH_2CH_2CH_2NH$, $NCH_2CH_2CH_2CH_2N$), 2.85-2.92 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.88-1.97 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.71-1.81 (m, 6H, $NCH_2CH_2CH_2NH$), 1.60-1.69 (m, 4H, $NCH_2CH_2CH_2CH_2N$)

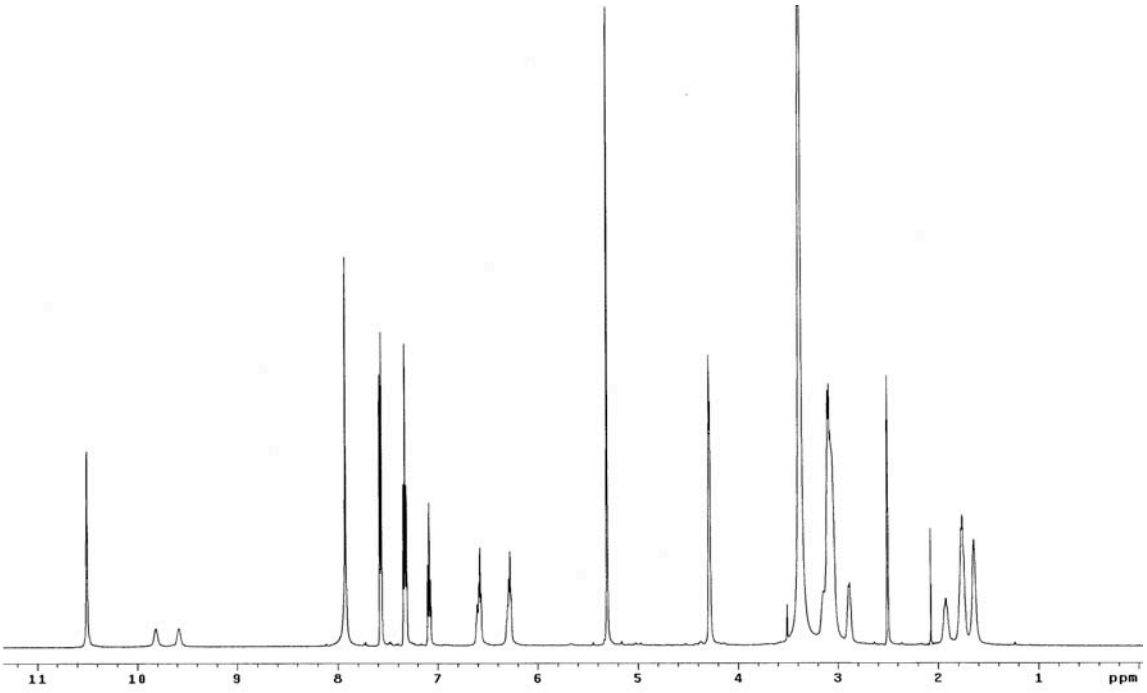
^{13}C NMR (75 MHz, DMSO-d6)

δ = 164.3, 158.3, 145.7, 138.4, 129, 124.2, 123.8, 119.2, 69.8, 52.1, 51.2, 51, 50, 49.9, 49, 36.5, 36.2, 35, 24.6, 21.5, 20.2

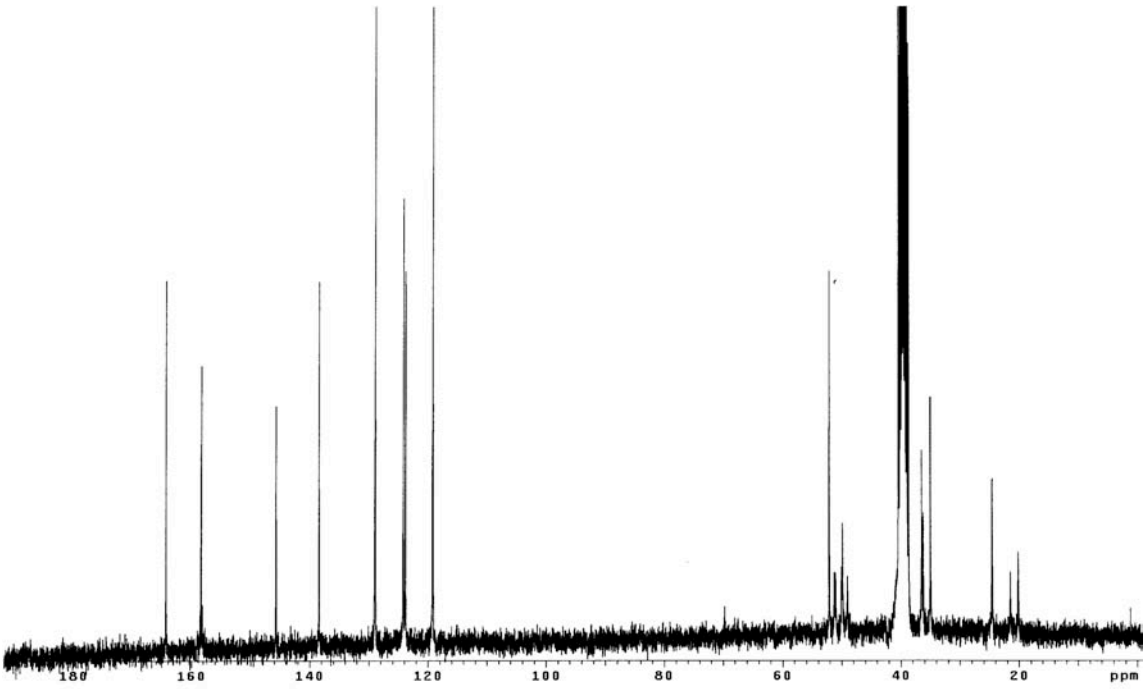
LRMS (MALDI) calc'd for $[C_{52}H_{73}N_{21}O_6 + Na]^+$ 1110.60, found 1110.6519

Analytical HPLC

homogeneous single peak, retention time = 14.1 min (5-95% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

10au:

^{1}H NMR (500 MHz, DMSO-d6)

δ = 10.42 (s, 3H, p-$CH_3C_6H_4N\mathit{H}CO$), 7.93 (bs, 2H, $NCH_2CH_2CH_2N\mathit{H}_2$), 7.92 (s, 3H), 7.45 (d, J=8.5, 6H), 7.12 (d, J=8.5, 6H), 6.56-6.63 (m, 3H, $N\mathit{H}CONHCH_2CH_2CH_2N$), 6.26-6.32 (m, 3H, $NHCON\mathit{H}CH_2CH_2CH_2N$), 5.28 (s, 6H), 4.26-4.29 (m, 6H), 3.0-3.17 (m, 18H, $NCH_2CH_2C\mathit{H}_2NH_2$, $NC\mathit{H}_2CH_2C\mathit{H}_2NH$, $NC\mathit{H}_2CH_2CH_2C\mathit{H}_2N$), 2.85-2.91 (m, 2H, $NC\mathit{H}_2CH_2CH_2NH_2$), 2.25 (s, 9H, p-$C\mathit{H}_3C_6H_4$), 1.88-1.97 (m, 2H, $NCH_2C\mathit{H}_2CH_2NH_2$), 1.70-1.80 (m, 6H, $NCH_2C\mathit{H}_2CH_2NH$), 1.61-1.69 (m, 4H, $NCH_2C\mathit{H}_2C\mathit{H}_2CH_2N$)

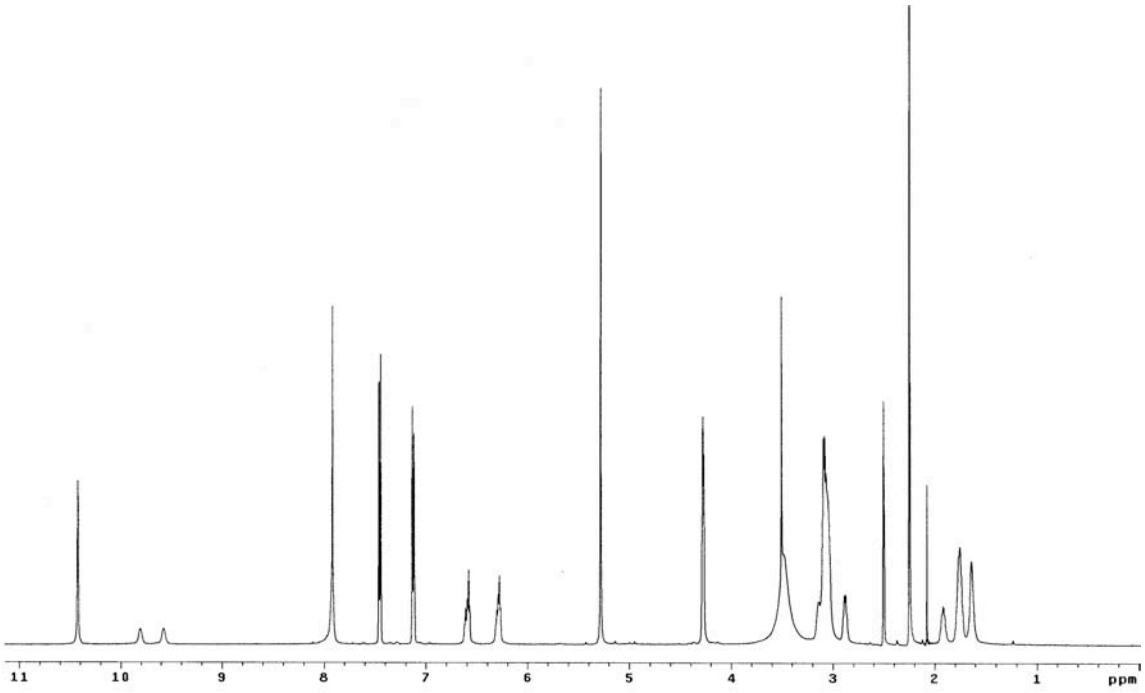
^{13}C NMR (125 MHz, DMSO-d6)

δ=164.1, 158.4, 158.3, 158.2, 145.7, 145.6, 135.9, 132.8, 129.3, 124.2, 119.2, 69.8, 52.1, 51.2, 51, 50, 49.8, 49, 36.4, 36.2, 35, 24.6, 21.5, 20.4, 20.1

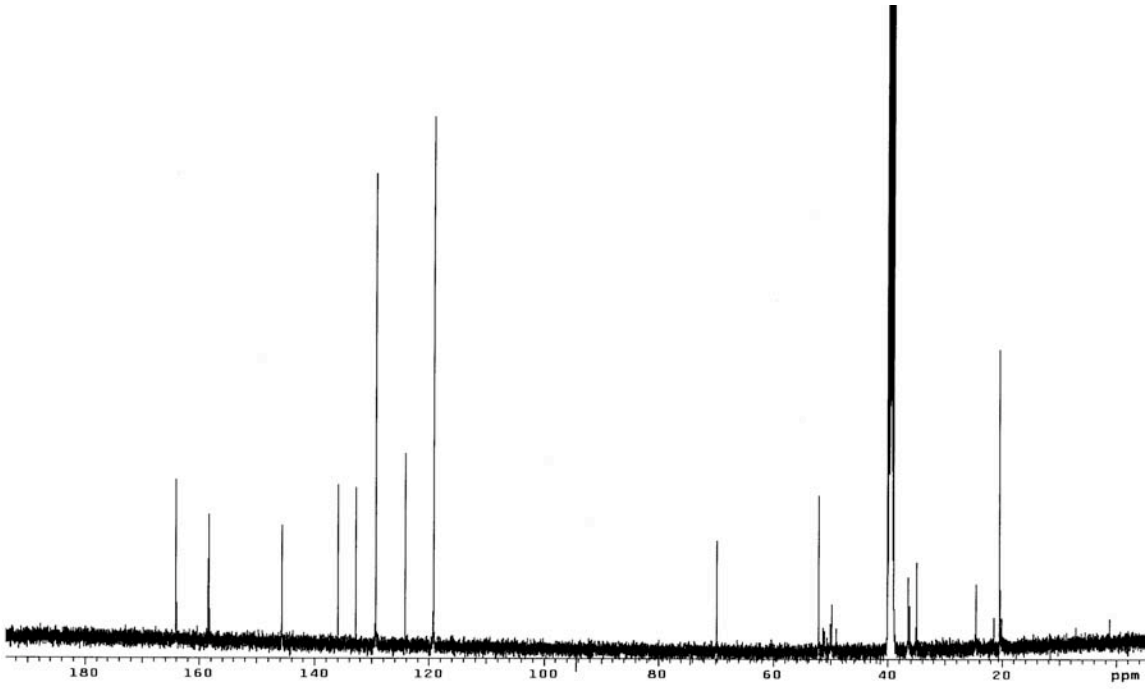
LRMS (MALDI) calc'd for $[C_{55}H_{79}N_{21}O_6 + H]^+$ 1130.66, found 1130.5867

Analytical HPLC

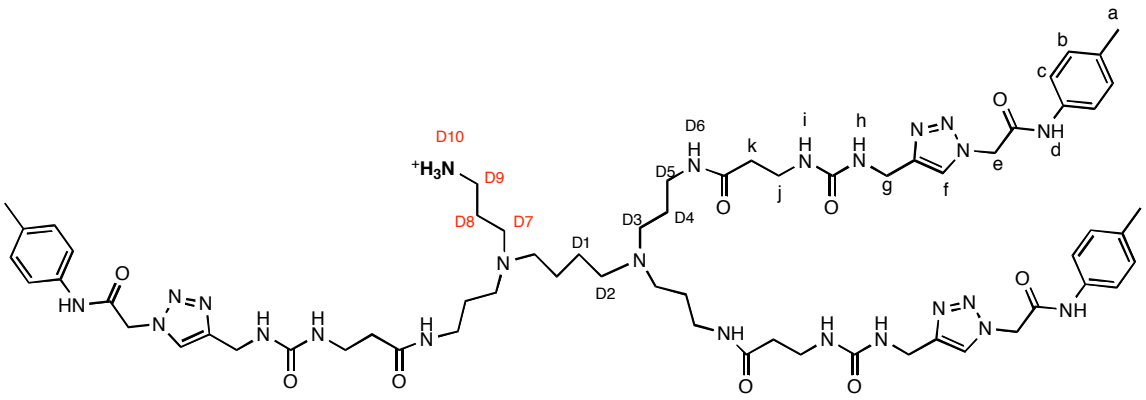
homogeneous single peak, retention time = 15.6 min (5-95% B in 30 min)



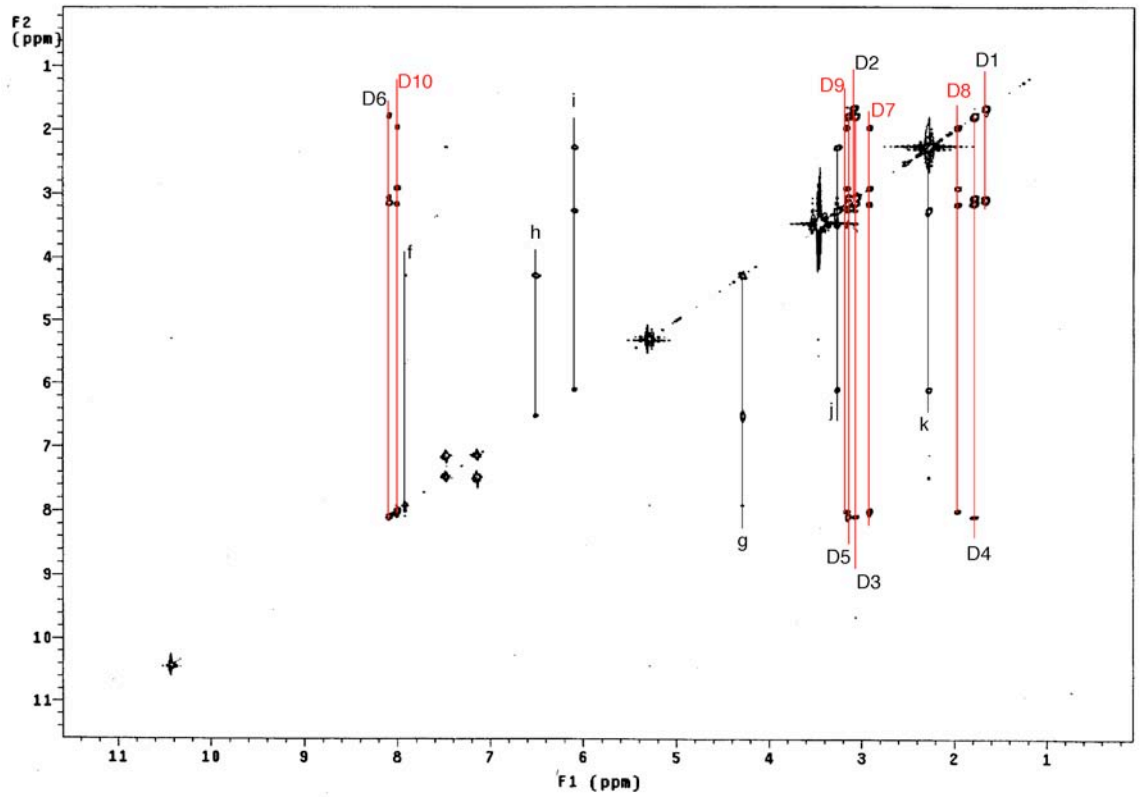
^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



10bu



80-ms TOCSY (DMSO-d6)

^{1}H NMR (500 MHz, DMSO-d6)

δ = 10.41 (s, 3H, p-$CH_3C_6H_4NHCO$), 8.05-8.10 (m, 3H, $NCH_2CH_2CH_2NHCO$), 7.99 (bs, 2H, $NCH_2CH_2CH_2NH_2$), 7.90 (s, 3H), 7.46 (d, J=8.0, 6H), 7.13 (d, J=8.0, 6H), 6.46-6.53 (m, 3H, $NHCONHCH_2CH_2CO$), 6.05-6.11 (m, 3H, $NHCONHCH_2CH_2CO$), 5.27 (s, 6H), 4.27 (d, J=4.35, 6H), 3.25 (q, J=4.55, 6H, $COCH_2CH_2NH$), 3.14 (m, 8H, $NCH_2CH_2CH_2NH_2$, $NCH_2CH_2CH_2NH$), 3.05 (m, 10H, $NCH_2CH_2CH_2CH_2N$, $NCH_2CH_2CH_2NH$), 2.86-2.93 (q, J=5.5, 2H, $NCH_2CH_2CH_2NH_2$), 2.26 (m, 15H, $COCH_2CH_2NH$, p-$CH_3C_6H_4$), 1.91-1.98 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.72-1.82 (m, 6H, $NCH_2CH_2CH_2NH$), 1.61-1.68 (m, 4H, $NCH_2CH_2CH_2CH_2N$)

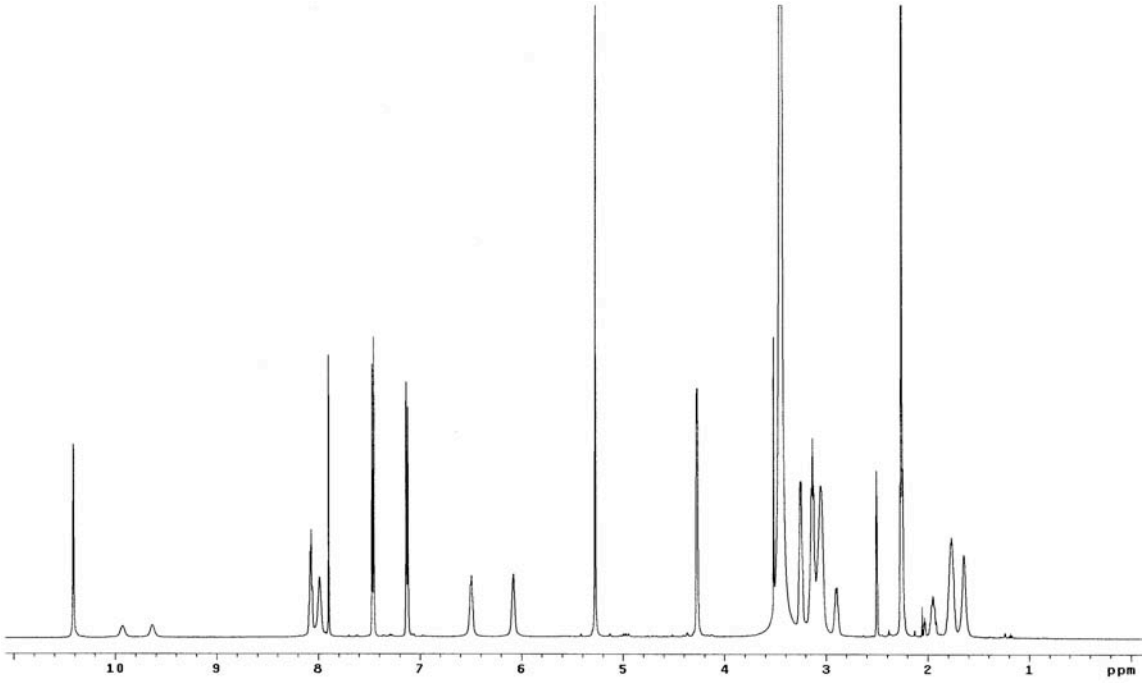
^{13}C NMR (125 MHz, DMSO-d6)

δ = 171.5, 164.2, 158.6, 158.4, 158, 145.9, 136, 133, 129.3, 124.2, 119.4, 118, 69.9, 52.2, 51.4, 51.3, 50.1, 50, 49.2, 36.5, 36.3, 36.1, 35.6, 35, 23.7, 21.6, 20.4, 20.3

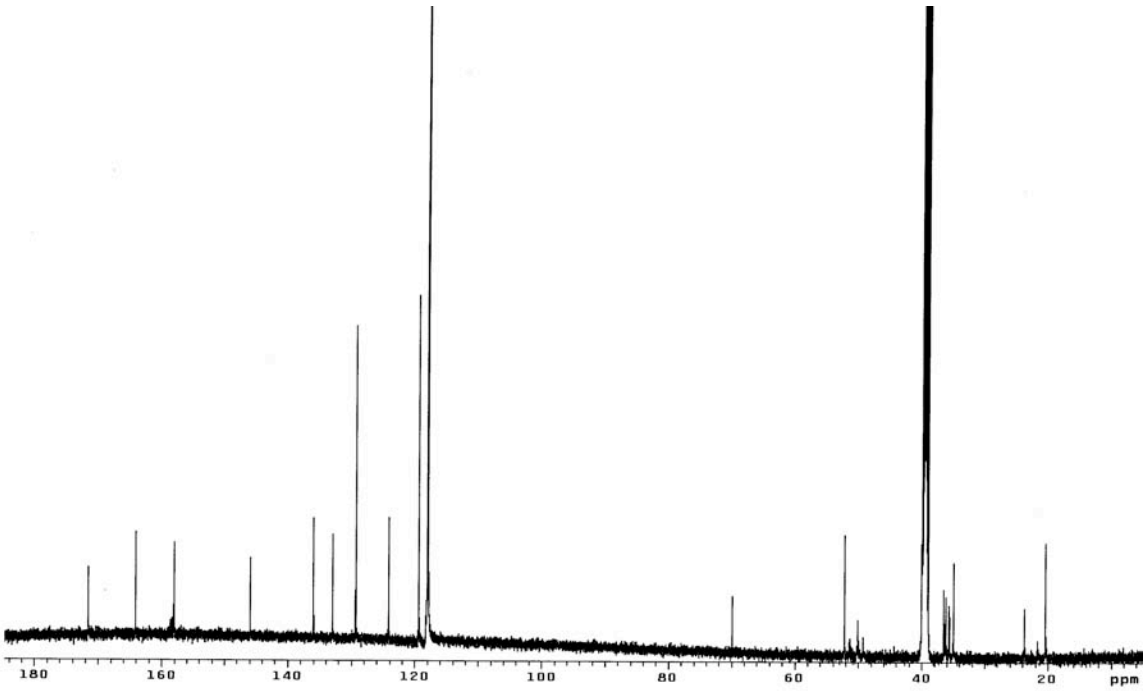
LRMS (MALDI) calc'd for $[C_{64}H_{94}N_{24}O_9 + H]^+$ 1344.59, found 1344.44

Analytical HPLC

homogeneous single peak, retention time = 15.7 min (5-95% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

10by:

^{1}H NMR (500 MHz, DMSO-d6)

δ = 9.71 (s, 3H, o,m-2$CH_3OC_6H_3NHCO$), 8.04-8.09 (m, 3H, $NCH_2CH_2CH_2NHCO$), 7.90 (s, 3H), 7.89 (bs, 2H, $NCH_2CH_2CH_2NH_2$), 7.66 (s, 3H), 6.99 (d, J=8.5, 3H), 6.66 (d, J=8.5, 3H), 6.43-6.49 (m, 3H, $NHCONHCH_2CH_2CO$), 6.02-6.07 (m, 3H, $NHCONHCH_2CH_2CO$), 5.41 (s, 6H), 4.25 (d, J=5.5, 6H), 3.81 (s, 9H), 3.66 (s, 9H), 3.23 (q, J=5.5, 6H, $COCH_2CH_2NH$), 3.12 (m, 8H, $NCH_2CH_2CH_2NH_2$, $NCH_2CH_2CH_2NH$), 3.05 (m, 10H, $NCH_2CH_2CH_2CH_2N$, $NCH_2CH_2CH_2NH$), 2.85-2.91 (m, 2H, $NCH_2CH_2CH_2NH_2$), 2.24 (t, J=6.5, 6H, $COCH_2CH_2NH$), 1.87-1.96 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.70-1.81 (m, 6H, $NCH_2CH_2CH_2NH$), 1.58-1.67 (m, 4H, $NCH_2CH_2CH_2CH_2N$)

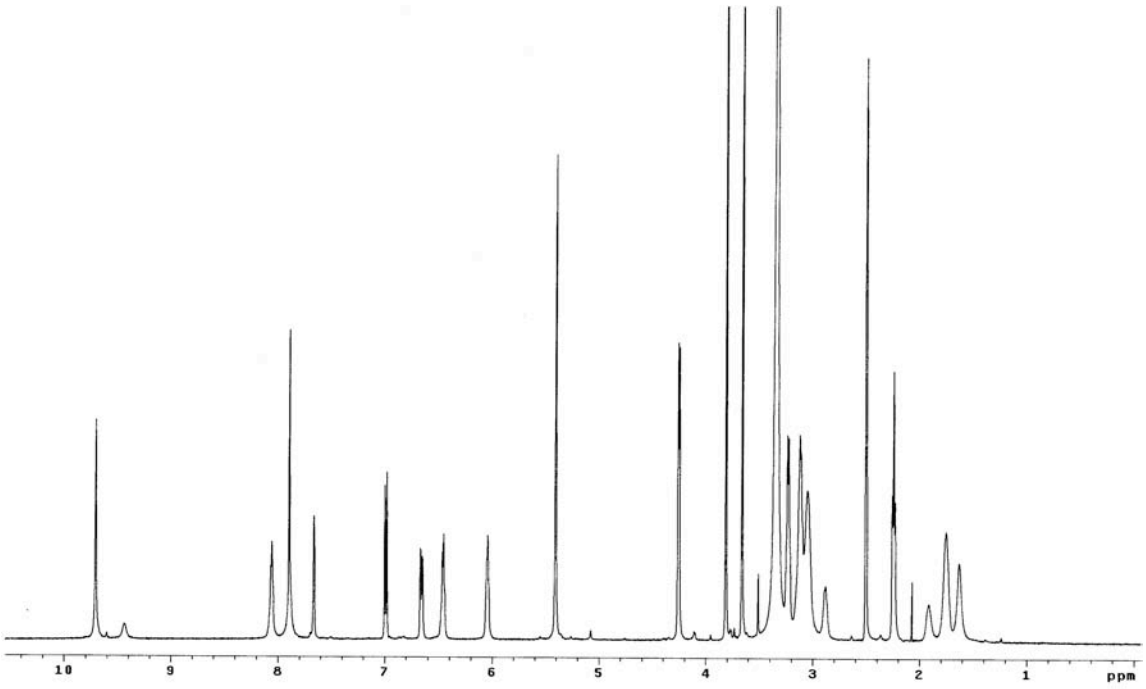
^{13}C NMR (75 MHz, DMSO-d6)

δ = 171.3, 164.8, 158.3, 157.9, 152.9, 145.8, 143.5, 128.9, 128.7, 128.4, 127.4, 124.2, 112, 108.6, 107.9, 69.8, 56.3, 55.3, 52.2, 51.3, 50, 49.1, 36.4, 36, 35.6, 34.9, 23.7, 21.5, 20.2

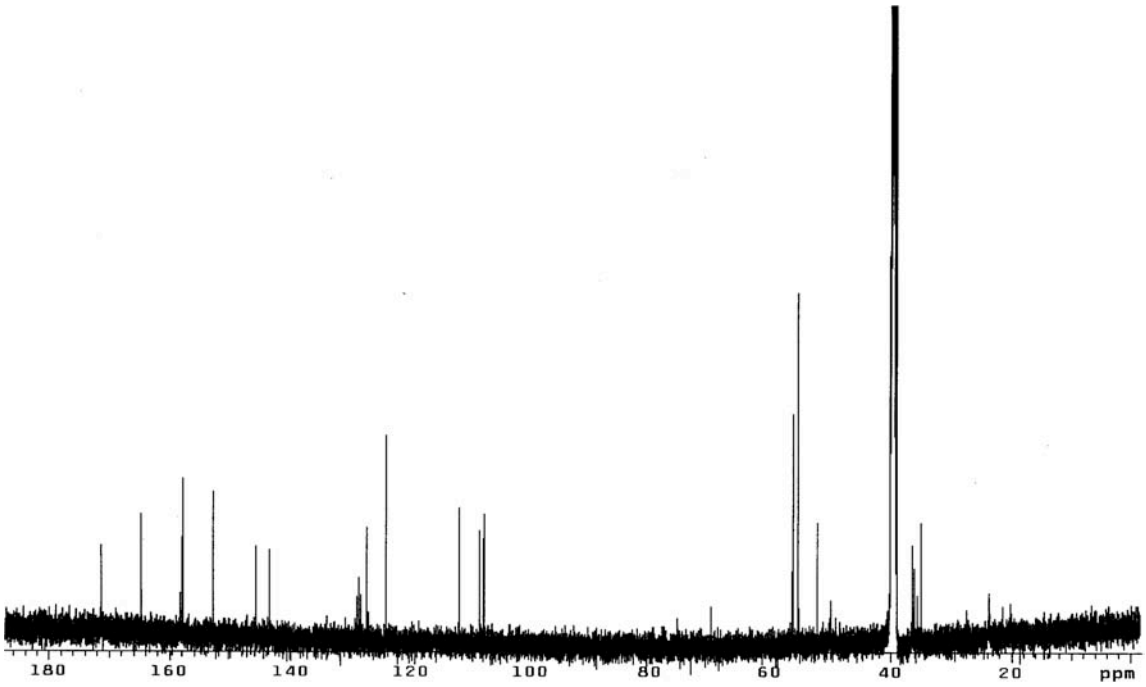
LRMS (MALDI) calc'd for $[C_{67}H_{100}N_{24}O_{15} + H]^+$ 1481.79, found 1481.915

Analytical HPLC

homogeneous single peak, retention time = 14.6 min (5-95% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

10bw:

^{1}H NMR (500 MHz, DMSO-d6)

δ = 9.73 (s, 3H, o,p-2$CH_3C_6H_3NHCO$), 8.07 (m, 3H, $NCH_2CH_2CH_2NHCO$), 7.91 (bs, 2H, $NCH_2CH_2CH_2NH_2$), 7.90 (s, 3H), 7.26 (d, J=8.0, 3H), 7.04 (s, 3H), 6.96 (d, J=8.0, 3H), 6.43-6.49 (m, 3H, $NHCONHCH_2CH_2CO$), 6.02-6.07 (m, 3H, $NHCONHCH_2CH_2CO$), 5.32 (s, 6H), 4.25 (d, J=5.5, 6H), 3.23 (q, J=5.5, 6H, $COCH_2CH_2NH$), 3.11 (m, 8H, $NCH_2CH_2CH_2NH_2$, $NCH_2CH_2CH_2NH$), 3.04 (m, 10H, $NCH_2CH_2CH_2CH_2N$, $NCH_2CH_2CH_2NH$), 2.84-2.91 (m, 2H, $NCH_2CH_2CH_2NH_2$), 2.22-2.26 (m, 15H, $COCH_2CH_2NH$, p-$CH_3C_6H_3$), 2.18 (s, 9H, o-$CH_3C_6H_3$), 1.87-1.95 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.70-1.80 (m, 6H, $NCH_2CH_2CH_2NH$), 1.58-1.67 (m, 4H, $NCH_2CH_2CH_2CH_2N$)

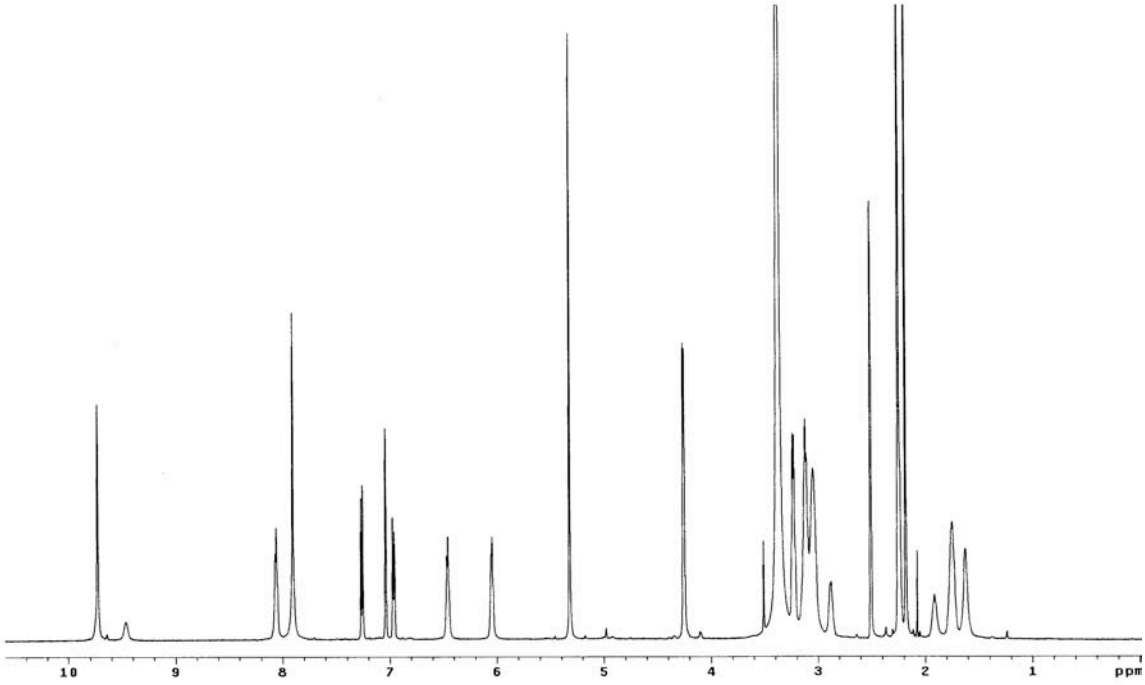
^{13}C NMR (75 MHz, DMSO-d6)

δ = 171.3, 164.5, 157.9, 145.8, 134.7, 132.9, 131.6, 131, 126.6, 124.8, 124.2, 69.8, 51.8, 51.4, 51.2, 50, 49.1, 36.4, 36.2, 36.1, 35.6, 34.9, 23.7, 20.5, 20.2, 17.7

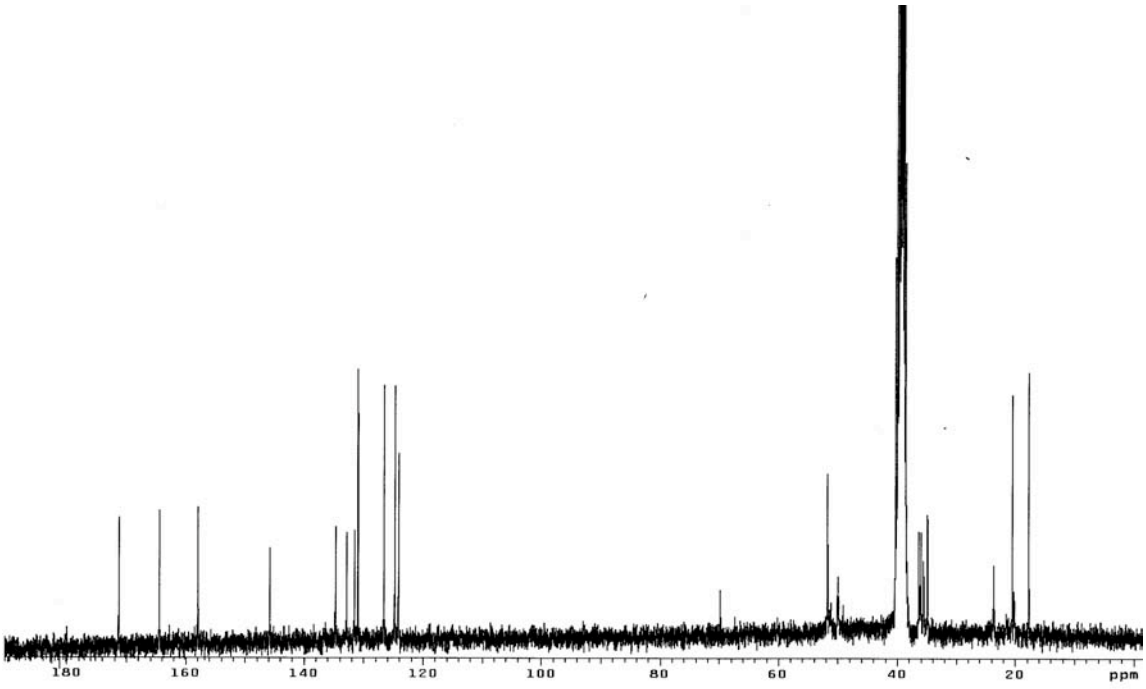
MS (MALDI) calc'd for $[C_{67}H_{100}N_{24}O_9 + H]^+$ 1385.82, found 1385.7897

Analytical HPLC

homogeneous single peak, retention time = 15.7 min (5-95% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

10cy:

^{1}H NMR (500 MHz, DMSO-d6)

δ = 9.72 (s, 3H, o,m-2$CH_3OC_6H_3NHCO$), 8.09-8.17 (m, 3H, $NCH_2CH_2CH_2NHCO$), 7.90 (bs, 2H, $NCH_2CH_2CH_2NH_2$), 7.82 (s, 3H), 7.66 (s, 3H), 7.23-7.28 (m, 6H), 7.16-7.21 (m, 9H), 6.99 (d, J=8.5, 3H), 6.65 (d, J=8.5, 3H), 6.52-6.59 (m, 3H, $NHCONHCH(CH_2C_6H_5)CO$), 6.24-6.31 (m, 3H, $NHCONHCH(CH_2C_6H_5)CO$), 5.40 (s, 6H), 4.26-4.33 (m, 3H, $NHCH(CH_2C_6H_5)CO$), 4.17-4.26 (m, 6H), 3.80 (s, 9H), 3.65 (s, 9H), 2.83-3.20 (m, 23H, $NCH_2CH_2CH_2NH_2$, $NCH_2CH_2CH_2NH$, $NCH_2CH_2CH_2CH_2N$, $NHCH(CH_2C_6H_5)CO$), 2.7-2.78 (m, 3H, $NHCH(CH_2C_6H_5)CO$), 1.87-1.95 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.66-1.78 (m, 6H, $NCH_2CH_2CH_2NH$), 1.57-1.66 (m, 4H, $NCH_2CH_2CH_2CH_2N$)

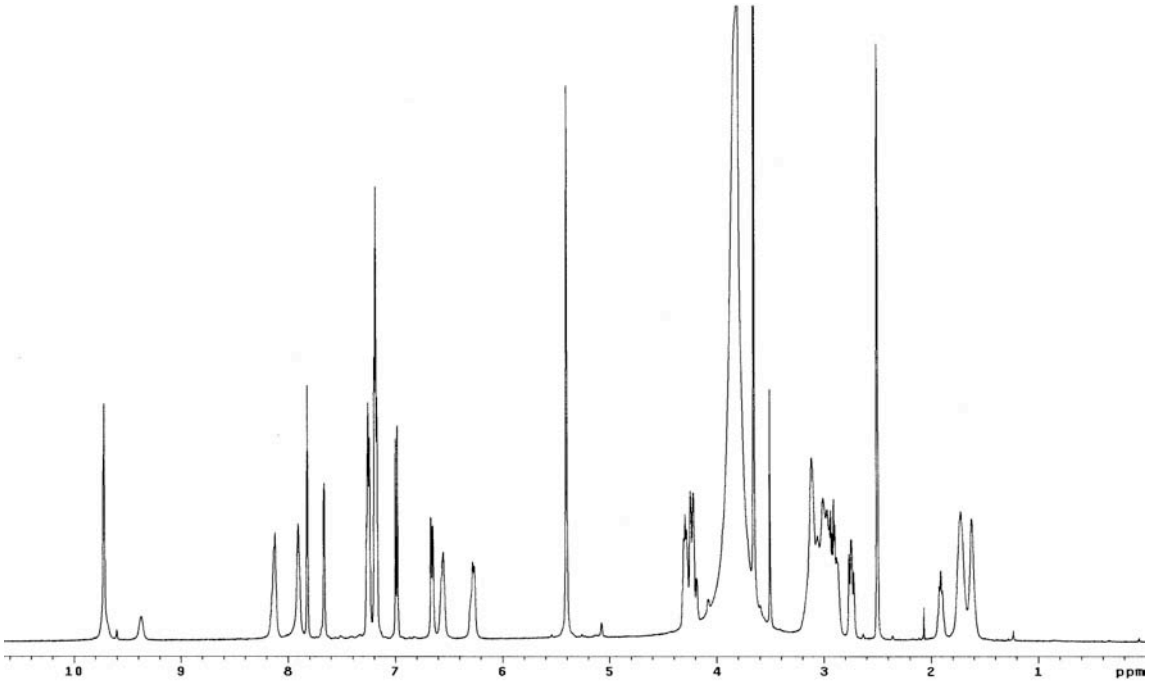
^{13}C NMR (75 MHz, DMSO-d6)

δ = 172.6, 164.7, 158.5, 158.1, 157.6, 157.4, 152.9, 145.6, 143.5, 137.8, 129.2, 128.2, 127.4, 126.3, 124.2, 112, 108.6, 107.9, 69.8, 56.3, 55.3, 54.9, 52.2, 36.3, 35.7, 34.9, 23.5, 21.5, 20.3

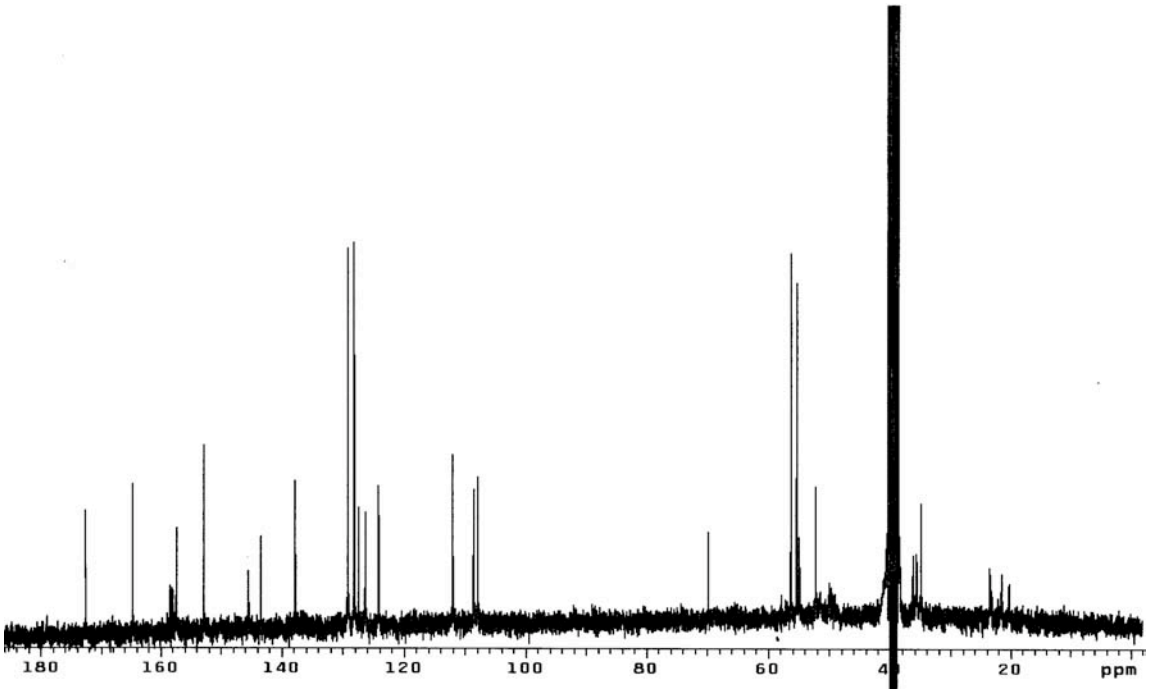
MS (MALDI) calc'd for $[C_{85}H_{112}N_{24}O_{15} + H]^+$ 1709.88, found 1710.0662

Analytical HPLC

homogeneous single peak, retention time = 15.6 min (5-95% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

10cz:

^{1}H NMR (500 MHz, DMSO-d6)

δ = 9.59 (s, 3H, o,m-2$CH_3OC_6H_3NHCO$), 8.09-8.17 (m, 3H, $NCH_2CH_2CH_2NHCO$), 7.89 (bs, 2H, $NCH_2CH_2CH_2NH_2$), 7.81 (s, 3H), 7.69 (d, J=8.5, 3H), 7.23-7.28 (m, 6H), 7.16-7.21 (m, 9H), 6.64 (s, 3H), 6.52-6.59 (m, 3H, $NHCONHCH(CH_2C_6H_5)CO$), 6.47 (d, J=8.5, 3H), 6.24-6.31 (m, 3H, $NHCONHCH(CH_2C_6H_5)CO$), 5.33 (s, 6H), 4.25-4.32 (m, 3H, $NHCH(CH_2C_6H_5)CO$), 4.17-4.25 (m, 6H), 3.83 (s, 9H), 3.73 (s, 9H), 2.83-3.20 (m, 23H, $NCH_2CH_2CH_2NH_2$, $NCH_2CH_2CH_2NH$, $NCH_2CH_2CH_2CH_2N$, $NHCH(CH_2C_6H_5)CO$), 2.7-2.78 (m, 3H, $NHCH(CH_2C_6H_5)CO$), 1.86-1.94 (m, 2H, $NCH_2CH_2CH_2NH_2$), 1.65-1.78 (m, 6H, $NCH_2CH_2CH_2NH$), 1.56-1.65 (m, 4H, $NCH_2CH_2CH_2CH_2N$)

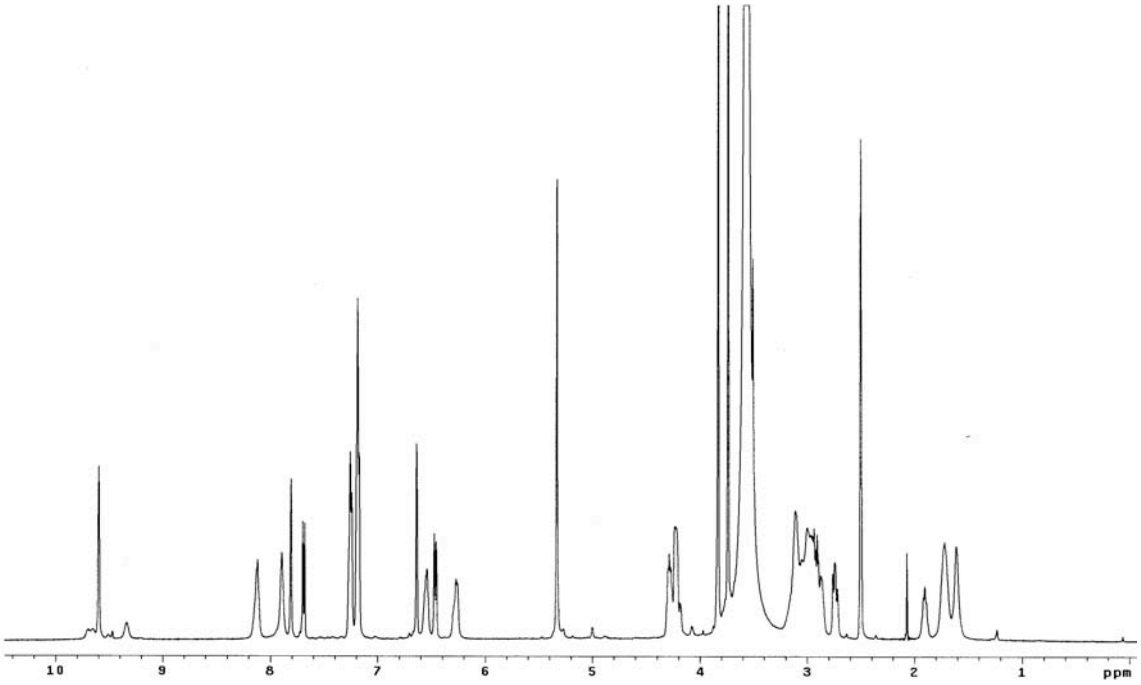
^{13}C NMR (75 MHz, DMSO-d6)

δ = 172.6, 164.2, 158.5, 158, 157.5, 157.4, 157, 151.4, 145.5, 145.3, 137.8, 129.2, 128.2, 126.4, 124.1, 123.3, 119.5, 104.1, 98.9, 69.8, 55.8, 55.3, 54.9, 52, 51.7, 51.4, 49.9, 49.2, 49, 36.3, 35.6, 34.9, 23.5, 23.4, 21.5, 20.3, 20.2

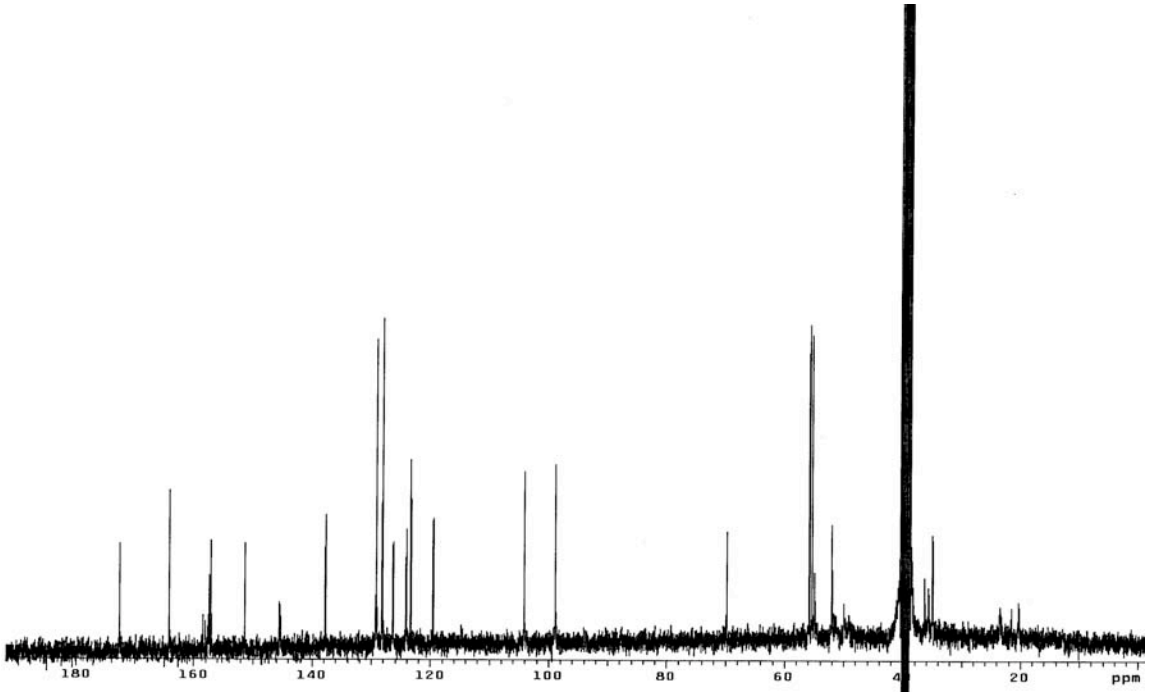
MS (MALDI) calc'd for $[C_{85}H_{112}N_{24}O_{15} + H]^+$ 1709.88, found 1710.4043

Analytical HPLC

homogeneous single peak, retention time = 15.2 min (5-95% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

10cx:

^{1}H NMR (500 MHz, DMSO-d6)

δ = 10.64 (s, 3H, p-$ClC_6H_4N\mathit{H}CO$), 8.13-8.20 (m, 3H, $NCH_2CH_2CH_2N\mathit{H}CO$), 7.90 (bs, 2H, $NCH_2CH_2CH_2N\mathit{H}_2$), 7.84 (s, 3H), 7.6 (d, J=8.5, 6H), 7.38 (d, J=8.5, 6H), 7.23-7.28 (m, 6H), 7.16-7.21 (m, 9H), 6.54-6.61 (m, 3H, $N\mathit{H}CONHCH(CH_2C_6H_5)CO$), 6.26-6.34 (m, 3H, $NHCON\mathit{H}CH(CH_2C_6H_5)CO$), 5.30 (s, 6H), 4.26-4.33 (m, 3H, $NHC\mathit{H}(CH_2C_6H_5)CO$), 4.17-4.26 (m, 6H), 2.82-3.19 (m, 23H, $NC\mathit{H}_2CH_2C\mathit{H}_2NH_2$, $NC\mathit{H}_2CH_2C\mathit{H}_2NH$, $NC\mathit{H}_2CH_2CH_2C\mathit{H}_2N$, $NHCH(C\mathit{H}_2C_6H_5)CO$), 2.7-2.77 (m, 3H, $NHCH(C\mathit{H}_2C_6H_5)CO$), 1.86-1.95 (m, 2H, $NCH_2C\mathit{H}_2CH_2NH_2$), 1.66-1.78 (m, 6H, $NCH_2C\mathit{H}_2CH_2NH$), 1.58-1.66 (m, 4H, $NCH_2C\mathit{H}_2C\mathit{H}_2CH_2N$)

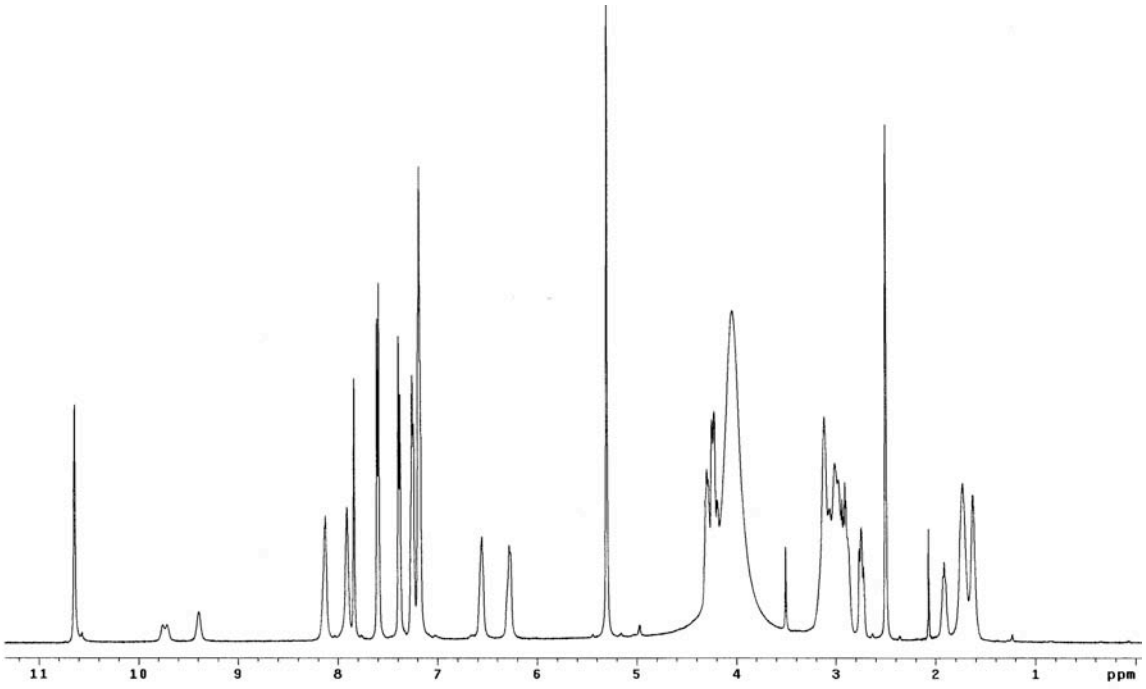
^{13}C NMR (75 MHz, DMSO-d6)

δ = 172.6, 164.5, 159, 158.6, 158.1, 157.5, 157.4, 145.5, 145.4, 137.8, 137.4, 129.2, 128.9, 128.2, 127.4, 126.3, 124.1, 120.8, 69.8, 54.9, 52.1, 51.5, 49.9, 49.1, 36.3, 35.6, 34.9, 23.5, 21.5, 20.2

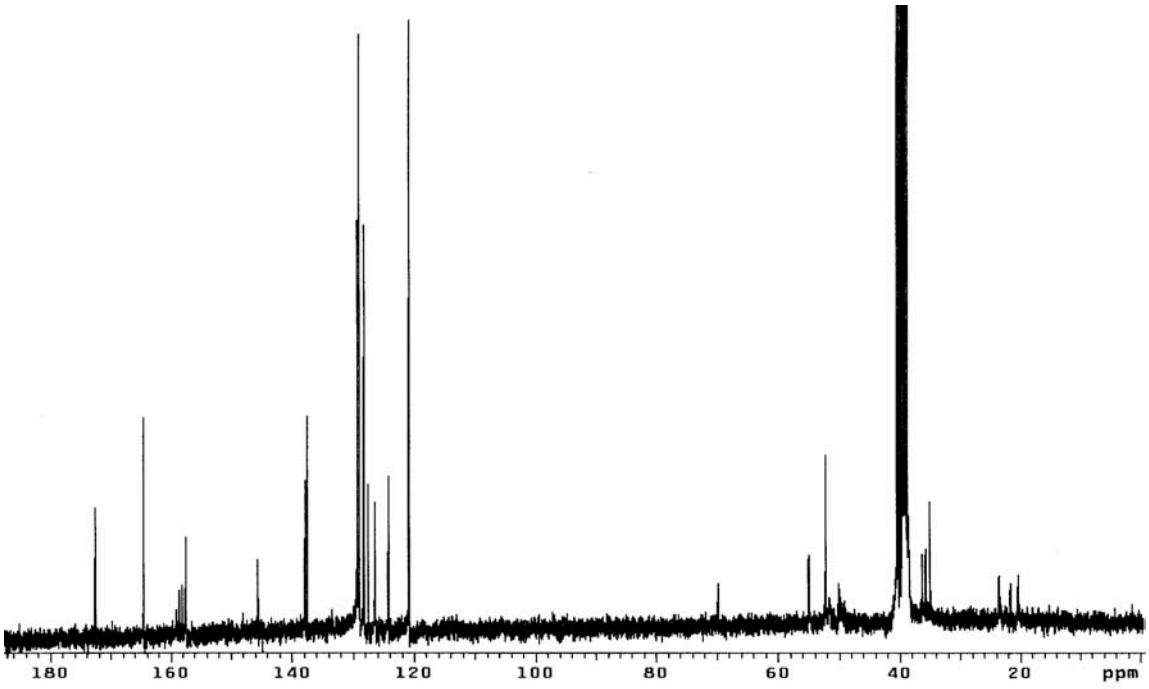
MS (MALDI) calc'd for $[C_{79}H_{98}Cl_3N_{24}O_9 + H]^+$ 1631.7, found 1631.1354

Analytical HPLC

homogeneous single peak, retention time = 19.5 min (5-95% B in 30 min)



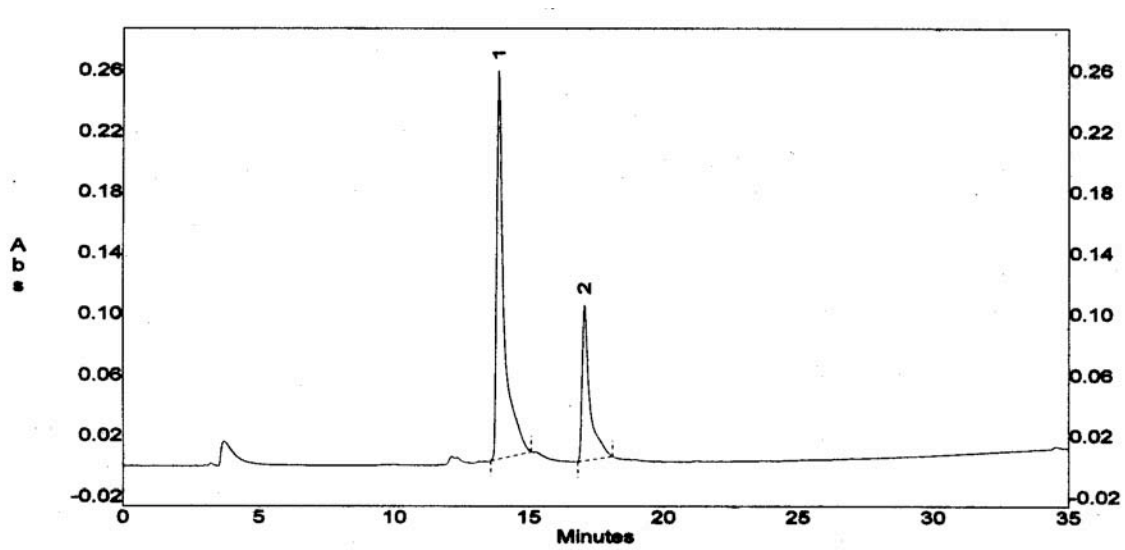
^{1}H NMR (DMSO-d6)



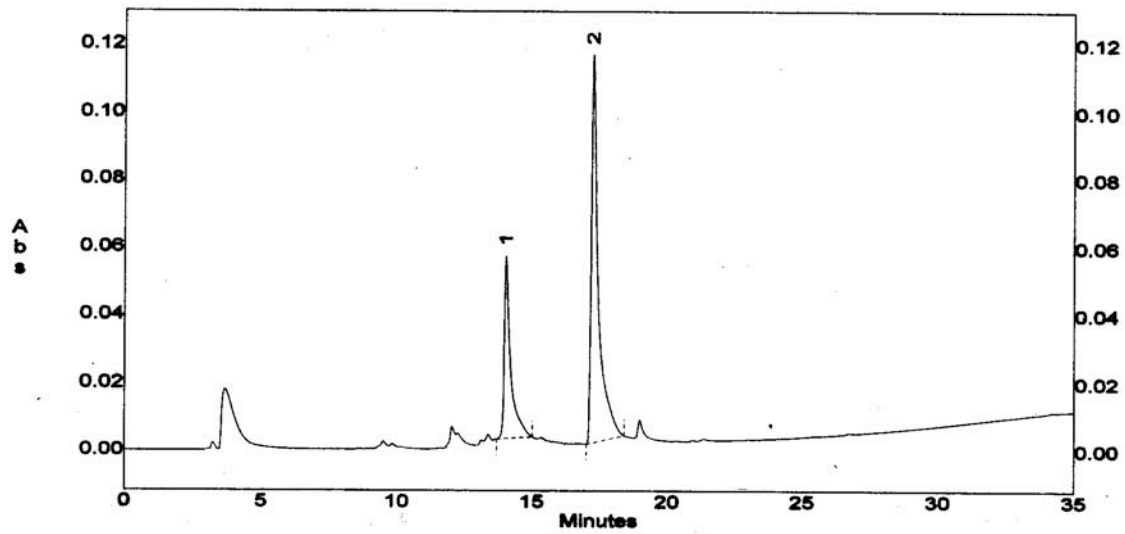
^{13}C NMR (DMSO-d6)

General Procedure for Immobilization of Multivalent Protein G Mimics

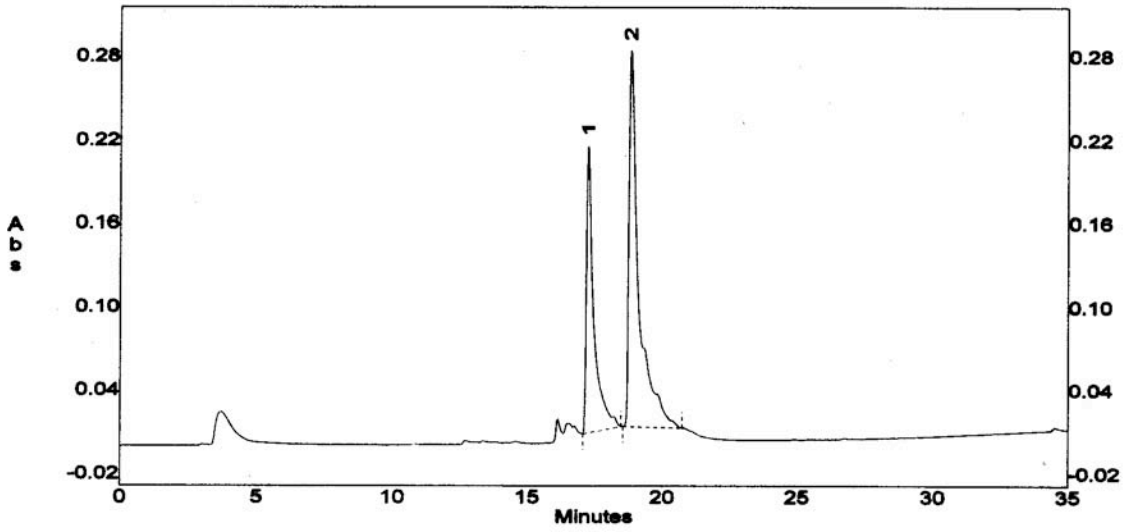
The solution of compounds **10**, 1,3-dimethoxybenzene, 0.2M $NaHCO_3$/DMF (1:1, pH = 9.7) was prepared and analyzed by RP-HPLC. Epoxy-activated sepharose 6B (Amersham Biosciences) was washed 10 times with pure water for 20 min. The mixture of the solution with sepharose was shaken for 20 h at 25 ˚C, then incubated in a water bath for 22 h at 35 ˚C. The reaction solution was analyzed again by RP-HPLC to monitor the absorption of compounds **10** immobilized on sepharose. The sepharose with the attachment of compounds **10** was washed three times with Tris (pH 9.4), total 10 h (2x3h and 1x4h) followed by pure water (5x).



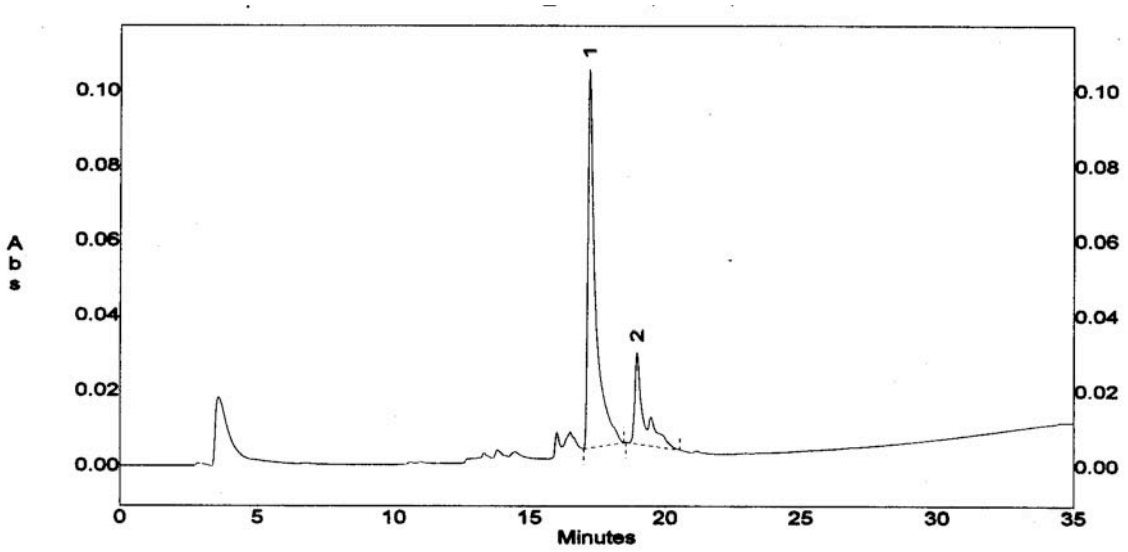
HPLC trace of the **10at (peak 1)** and **standard (peak 2)** solution in 0.2M $NaHCO_3$/DMF (1:1) at t = 0.



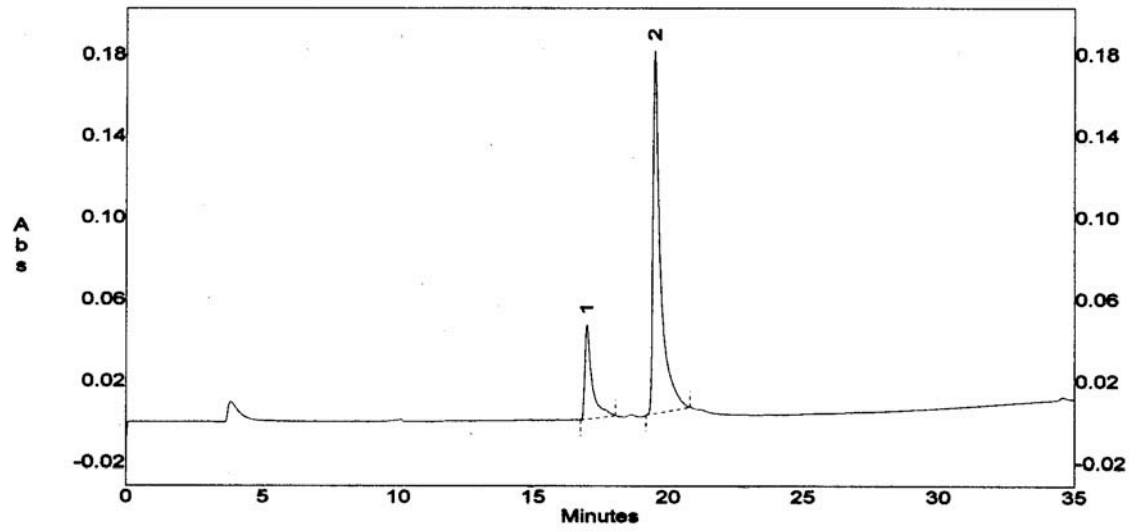
HPLC trace of the **10at (peak 1)** and the internal standard **(peak 2)** solution in 0.2M $NaHCO_3$/DMF (1:1) at the end



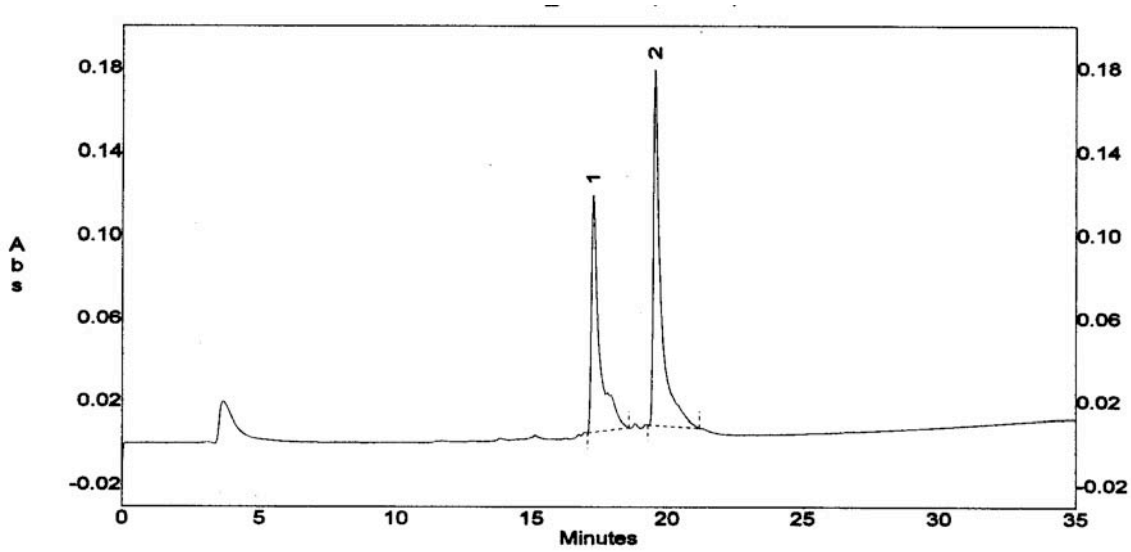
HPLC trace of the **10cw (peak 2)** and the internal standard**(peak 1)** solution in 0.2M $NaHCO_3$/DMF (1:1) at t = 0.



HPLC trace of the **10cw (peak 2)** and the internal standard **(peak1)** solution in 0.2M $NaHCO_3$/DMF (1:1) at the end of loading



HPLC trace of the **10cx (peak 2)** and the internal standard**(peak 1)** solution in 0.2M $NaHCO_3$/DMF (1:1) at t = 0.

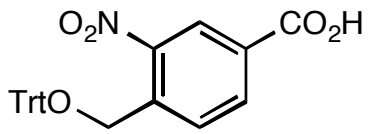


HPLC trace of the **10cx (peak 2)** and **standard (peak1)** solution in 0.2M $NaHCO_3$/DMF (1:1) at the end of loading

Docking Protocol

Docking/Affinity in Insight II was used for the docking study, and the Monte Carlo method was applied to search for the globally optimized complexes. Crystal structure of Fab fragment of IgG was selected as the receptor for this affinity modeling. Receptor and ligands were prepared using Insight II. Domain III of protein G and water molecules were removed from the complex. Remaining Fab fragment of IgG was used as a receptor. Ligands **A-C** (Figure 4.5) were constructed and optimized using Builder module of Insight II. A subset containing the ligand and the specified binding active site was then defined. CVFF force field was selected to assign potentials for the ligand-receptor assembly. Affinity and electrostatic grids were generated by Affinity/GridDocking module. The partial charges on the atoms of IgG/Fab were taken from the CVFF force field parameters. The grid spacing was 0.5 Å in each dimension, and each grid map consisted of 70 x 70 x 70 grid points. The subset was minimized with 2000 steps using Conjugate Gradient (Polak_Ribiere) method, and energetically favorable structures were found via Monte Carlo search at the same time.

APPENDIX G
EXPERIMENTAL FOR CHAPTER V



39

Procedure

A mixture of 4-(bromomethyl)-3-nitrobenzoic acid (5.0 g, 19.23 mmol) and sodium carbonate (8.15 g, 76.91 mmol) in 84 mL of 1:1 acetone/water was refluxed. After 3 h the mixture was cooled to 25 ˚C, then acidified to pH 2 with 2M HCl, extracted with ethyl acetate (80 mL x 4), dried over anhydrous magnesium sulfate and concentrated *in vacuo* to give 4-(Hydroxymethyl)-3-nitrobenzoic acid 3.5 g (92%) as a brown solid. This material was used for the next step without further purification.

4-(Hydroxymethyl)-3-nitrobenzoic acid (3.5 g, 17.77 mmol), triphenylmethyl chloride (5.94 g, 21.33 mmol), silver nitrate (3.62 g, 21.33 mmol), *N*,*N*-dimethylformamide (45 mL) and diisopropylethylamine (6.8 mL, 39.09 mmol) were added to a 100 mL round bottom flask and the mixture was stirred at 55 ˚C for 10 h. The solvent was removed *in vacuo* and the residue was purified by flash chromatography with an eluent of 20 - 50% ethyl acetate and 1% acetic acid in hexane to give 5.125 g (yield 66%) compound **39** as yellow solid.

Mp 218 - 220 °C

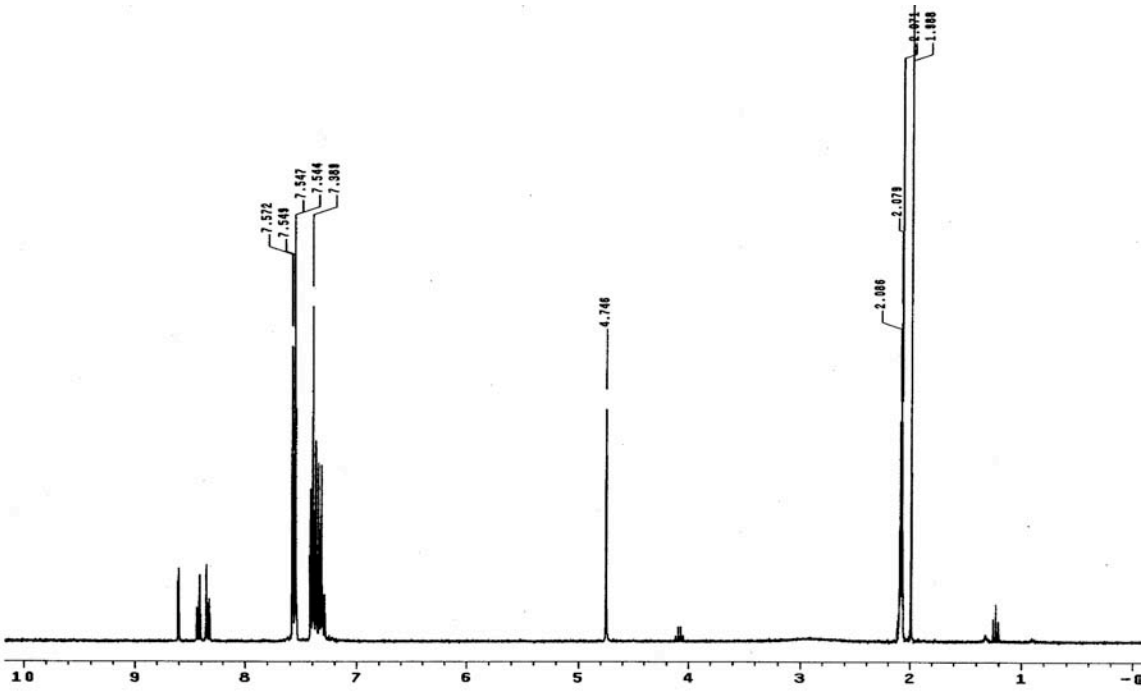
$\boldsymbol{R_f}$ = 0.44 (1:2:0.1 ethyl acetate/hexane/AcOH)

^{1}H NMR (300 MHz, acetone-d6)

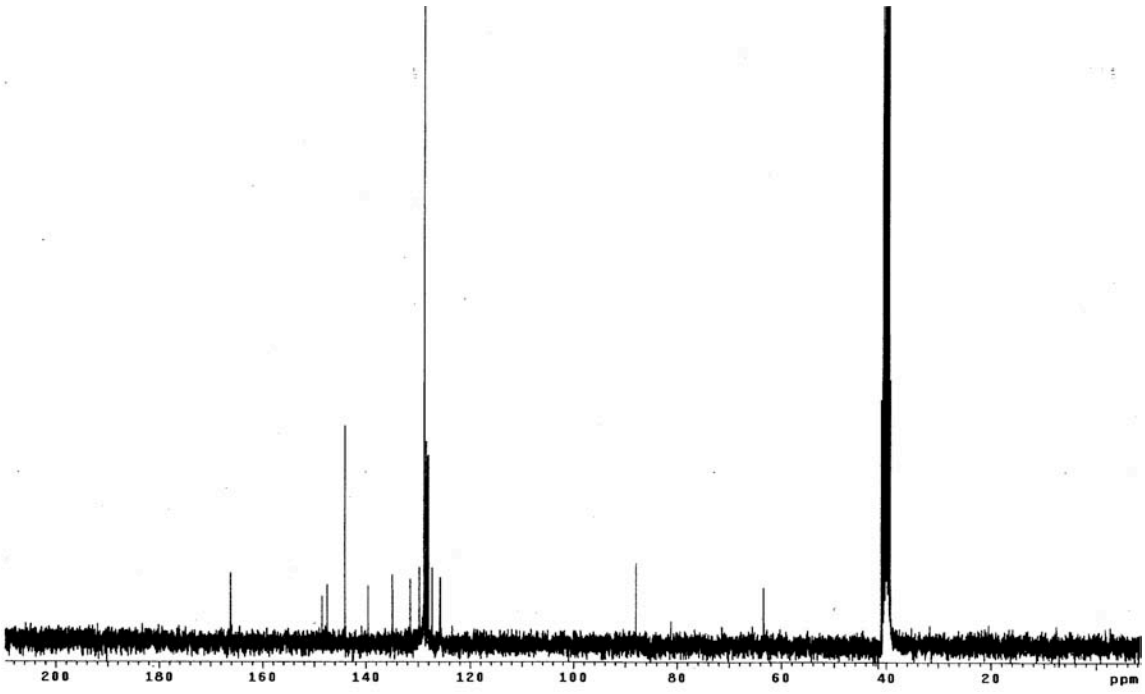
 = 8.61 (d, *J*=2.3, 1H), 8.43 (dd, J=6.2 and 2.3, 1H), 8.33 (d, J=6.2, 1H), 7.54 - 7.59 (m, 6H), 7.29 - 7.42 (m, 9H), 4.75 (s, 2H)

^{13}C NMR (75 MHz, DMSO-d6)

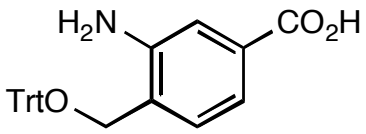
 = 166.1, 148.4, 147.4, 144.0, 139.5, 134.8, 131.5, 129.8, 128.8, 128.4, 128.2, 128.0, 127.3, 125.8, 88.0, 63.4



^{1}H NMR (acetone-d6)



^{13}C NMR (DMSO-d6)



40

Procedure

To a solution of 5.125 g (11.67 mmol) **39** in 3:1 ethyl acetate/methanol (120 mL) was added PtO_2 (0.35 g, 1.54 mmol). The mixture was stirred vigorously under H_2 at 25 ˚C for 5 h. The reaction mixture was filtered through celite and concentrated to dryness *in vacuo*. The residue was purified by flash chromatography with 30% EtOAc : 70% hexane to give 3.605 g (76%) of the product **40** as pale yellow solid.

Mp 180 - 182 °C

R_f = 0.42 (30% ethyl acetate in hexane)

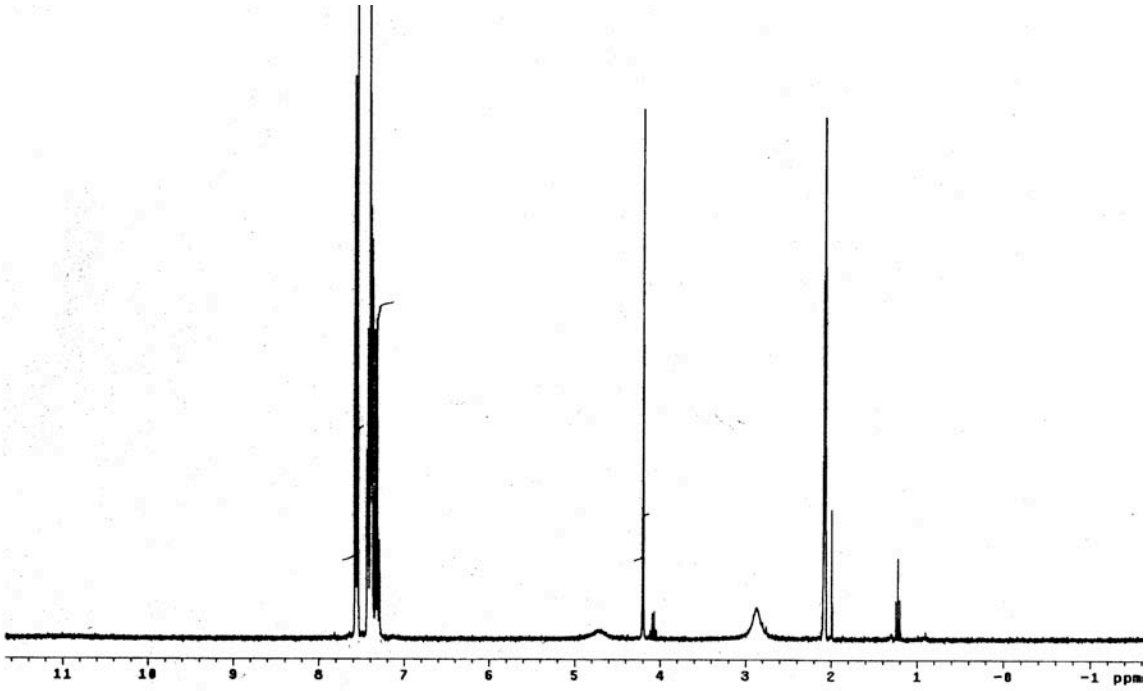
^{1}H NMR (300 MHz, acetone-d6**)**

 = 7.55 - 7.59 (m, 6H), 7.30 - 7.45 (m, 12H), 4.2 (s, 2H)

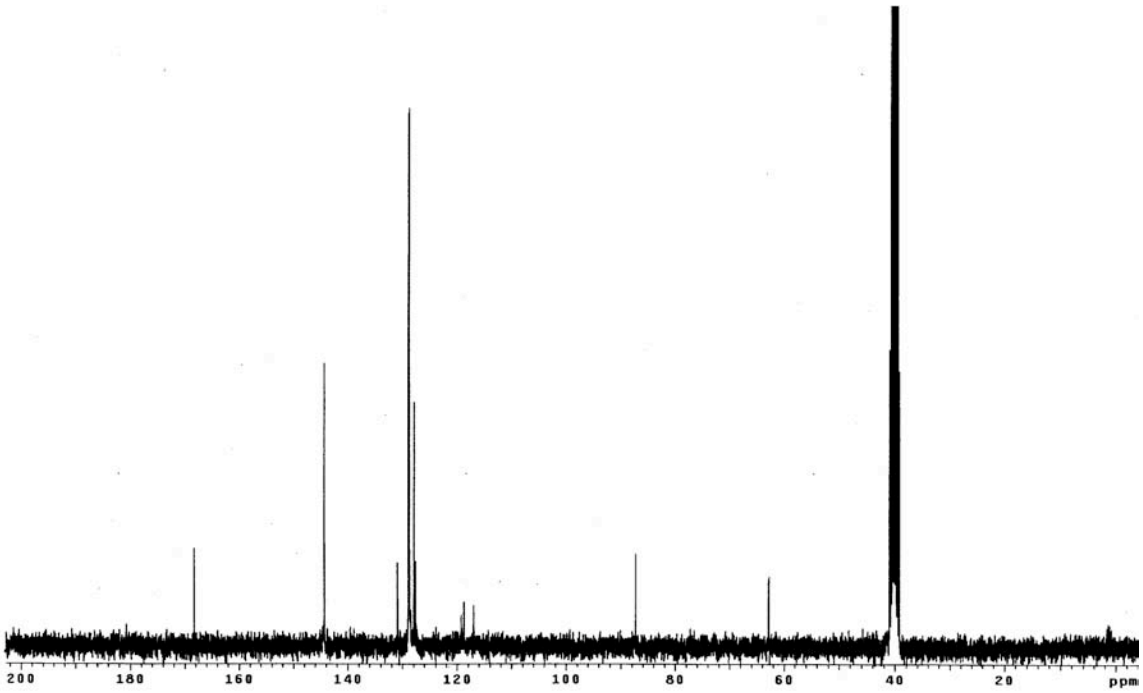
^{13}C NMR (75 MHz, DMSO-d6)

 = 168.2, 144.3, 130.9, 128.9, 128.8, 127.9, 127.6, 118.7, 117.0, 87.3, 62.8

MS (ESI) calc'd for $C_{27}H_{23}NO_3$ 409.2, found 409.2



^{1}H NMR (acetone-d6)



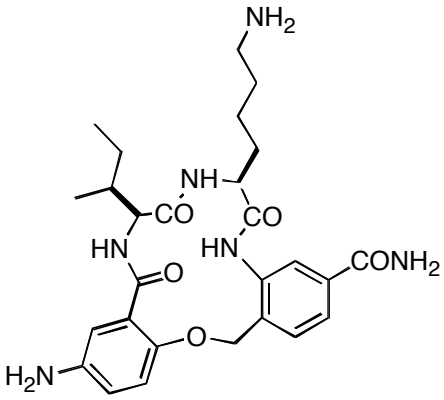
^{13}C NMR (DMSO-d6)

Procedure for Coupling Compound 40 to Polystyrene-Rink Resin

Resin with Rink linker was swelled in CH_2Cl_2 in a fritted syringe for 30 min. The Fmoc protecting group on the resin was removed by treating the resin with 20% piperidine in DMF (2 x, 10□min and then 15 min). The resin was washed with DMF (3x), MeOH (3x), CH_2Cl_2 (3x), DMF (2x), MeOH (2x), and CH_2Cl_2 (3x) and then treated with template **40** (2 equiv), HBTU (3 equiv), HOBt (3 equiv), and DIEA (8 equiv) in DMF. After 2 h of gentle shaking, the mixture was drained. The resin was then washed as the above washing cycle. A ninhydrin test on a small sample of beads was performed and gave a negative result.

General Procedure for Preparation of the Cyclic Peptidomimetics (36a-d,NH_2) Illustrated for Compound 36a. The resin containing template **40** was treated with Fmoc-Lys(Boc)-OH (4 equiv), PyBrop (4.6 equiv), 2,6-lutidine (15 equiv) in CH_2Cl_2 for 20 h. After washing with DMF (3x), MeOH (3x), CH_2Cl_2 (3x), DMF (2x), MeOH (2x), and CH_2Cl_2 (3x), and Fmoc deprotection, Fmoc-Ile was then introduced using Fmoc-Ile-OH (3 equiv), DIC (3 equiv), HOBt (3 equiv), and DIEA (5 equiv) in CH_2Cl_2/DMF 2:1 for 2 h. The previously described washing cycle and Fmoc deprotection were repeated and 2-fluoro-5-nitrobenzoic acid moiety was introduced by treating the resin with 2-fluoro-5-nitrobenzoyl chloride (2.5 equiv) and DIEA (5 equiv) in CH_2Cl_2 for 1 h. The side-chain protecting group (Trt) of the template was removed by treatment with 1% TFA and 5% TIS in CH_2Cl_2 (6 x 3 min). After the resin was washed and the macrocyclization step was carried out by gently shaking the resin supported peptide with K_2CO_3 (10 equiv) in DMF at 25 ˚C for 40 h, The mixture was drained and the resin was washed with H_2O (5x), DMF (3x), MeOH (3x), DMF (2x), MeOH (2x), CH_2Cl_2 (2x), MeOH (2x), and CH_2Cl_2 (3x) and then dried *in vacuo* over night. The resin containing side-chain protected cyclic peptide was swelled in CH_2Cl_2 for 30 min and treated with $SnCl_2.2H_2O$ (10 equiv) in DMF for 24 h. After washing the peptide was cleaved from the resin by the treatment of a mixture of 90% TFA, 5% TIS, and 5% H_2O for 2 h. The cleavage solution was separated from the resin via filtration. After most of the cleavage

cocktail was evaporated in a stream of nitrogen, the crude peptide was precipitated using cooled anhydrous ethyl ether. The purity of this crude product was determined by analytical HPLC (SSI system, 2-40% B in 30 min). The crude product was dissolved in H_2O, filtered, then purified with preparative HPLC and finally lyophilized to yield white powder of **36a:** 15 mg, yield: 40%.



36a

^{1}H NMR (500 MHz, DMSO-d6)

δ = 8.79 (s, 1H), 8.62 (d, *J* = 8.28 Hz, 1H), 8.44 (d, *J* = 9.11 Hz, 1H), 8.40 (s, 1H), 7.98 (s, 1H), 7.75-7.63 (m, 4H), 7.61 (dd, *J* = 7.97 Hz, 1.71 Hz, 2H), 7.43 (t, *J* = 8.0, 2H), 7.26 (s, 1H), 7.19 (d, *J* = 8.5 Hz, 1H), 5.45 (d, *J* = 12.3 Hz, 1H), 5.27 (d, *J* = 12.3 Hz, 1H), 4.41 (m, 1H), 4.26 (t, *J* = 9.5 Hz, 1H), 2.81-2.70 (m, 2H), 1.94 (m, 2H), 1.74 (m, 1H), 1.57 (m, 3H), 1.40 (m, 2H), 1.21 (m, 1H), 0.93 (d, *J* = 7.5 Hz, 3H), 0.89 (d, *J* = 7.5 Hz, 3H)

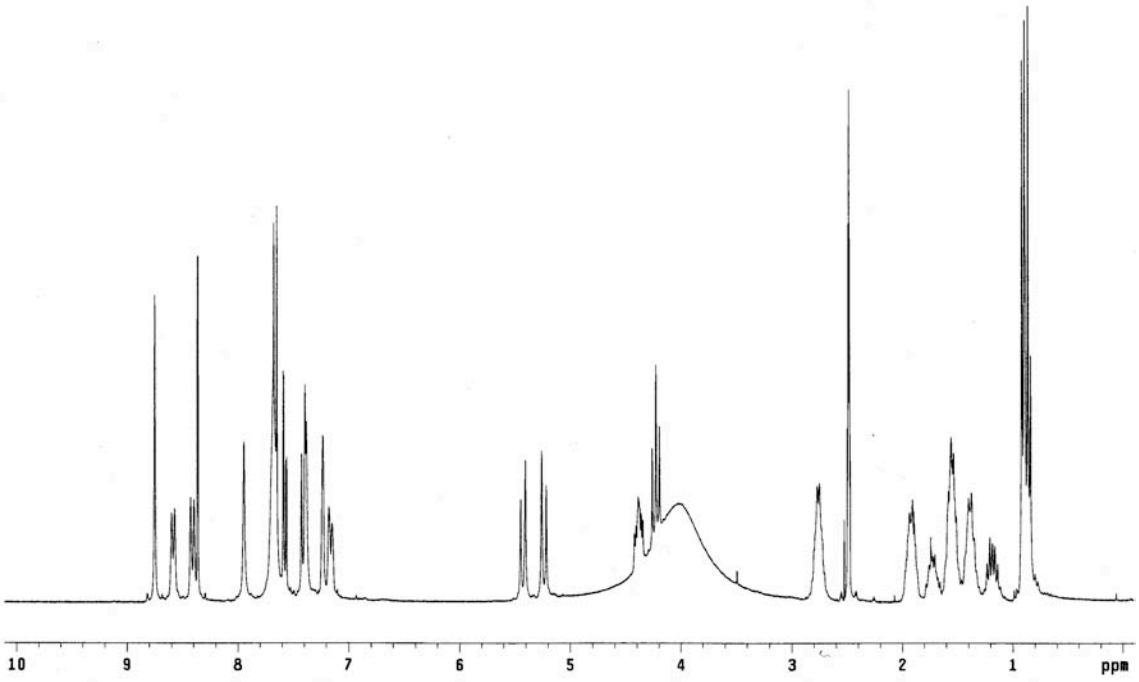
^{13}C NMR (75 MHz, DMSO-d6)

δ = 172.3, 170.9, 167.9, 165.9, 137.7, 136.0, 131.3, 129.8, 127.0, 123.5, 122.7, 115.6, 68.4, 60.7, 53.7, 35.7, 31.0, 27.1, 26.0, 23.3, 16.3, 11.3

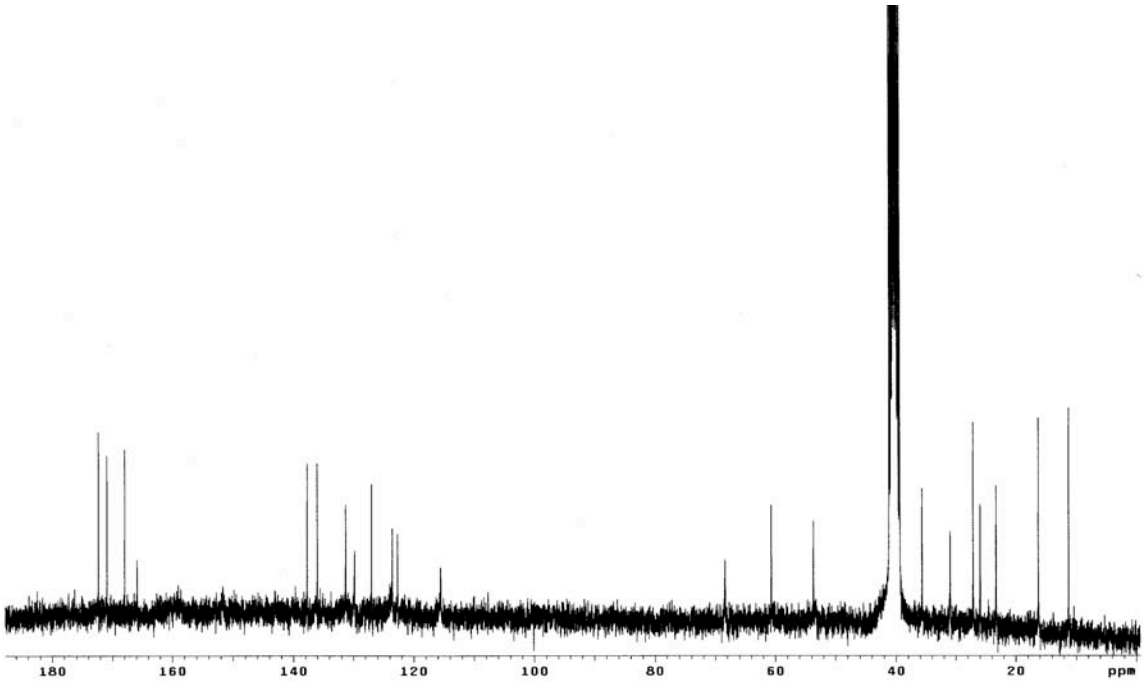
MS (MALDI) calc'd for $C_{27}H_{36}N_6O_5$ ($M+H^+$) 525.6, found 525.1

Analytical HPLC

homogeneous single peak, retention time = 11.7 min (2-40% B in 30 min)

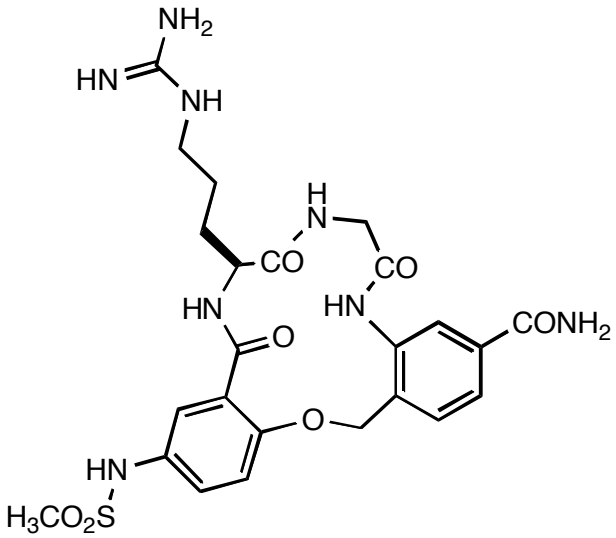


^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

General Procedure for Preparation of the Derivatized Macrocycles 37 and 38 *N*-Methyl Sulfonamide Derivatives 4a-e. Resin containing side-chain protected cyclic peptide with amino group was swelled in CH_2Cl_2 in a fritted syringe for 30 min and treated with 10 equiv. of methanesulfonyl chloride, 10 equiv. of pyridine and 1 equiv. of 4-dimethylaminopyridine in CH_2Cl_2. After gental shaking for 24 h, the resin was washed with DMF (3x), MeOH (3x), CH_2Cl_2 (3x), DMF (2x), MeOH (2x), and CH_2Cl_2 (3x) and the peptide was cleaved from the resin by treatment with a mixture of 90% TFA, 5% TIS, and 5% H_2O for 2 h. The cleavage solution was separated from the resin via filtration. After most of the cleavage cocktail (about 90%) was evaporated in a stream of nitrogen, the crude peptide was triturated using anhydrous ethyl ether. The purity of this crude product was determined by analytical HPLC (SSI system, 2-40% B in 30 min). The crude product was dissolved in H_2O, filtered, then purified with preparative HPLC and finallly lyophilized to yield products **37a-e**.



37d

^{1}H NMR (300 MHz, DMSO-d6)

 = 9.57 (s, 1H), 8.79 (d, *J*=8.4 Hz, 1H), 8.68 (s, 1H), 8.61 (t, *J*=6.3, 1H), 8.33 (d, *J* = 1.86 Hz, 1H), 7.99 (s, 1H), 7.66 (dd, *J*=7.8 Hz, 1.59 Hz, 2H), 7.56 (t, *J*=5.1 Hz, 2H), 7.39 (m, 3H), 7.24 (m, 3H), 5.33 (d, *J* = 12.6 Hz, 1H), 5.19 (d, *J* = 12.6 Hz, 1H), 4.42 (m, *J*=7.5 Hz, 1H), (m, 1H), 4.0 (dd, *J* = 16.8 Hz, 6.0 Hz, 1H), 3.86 (dd, *J* = 16.8 Hz, 6.0 Hz, 1H), 3.15 (m, 2H), 2.94 (s, 3H), 1.79 (m, 1H), 1.68 (m, 1H), 1.55 (m, 2H)

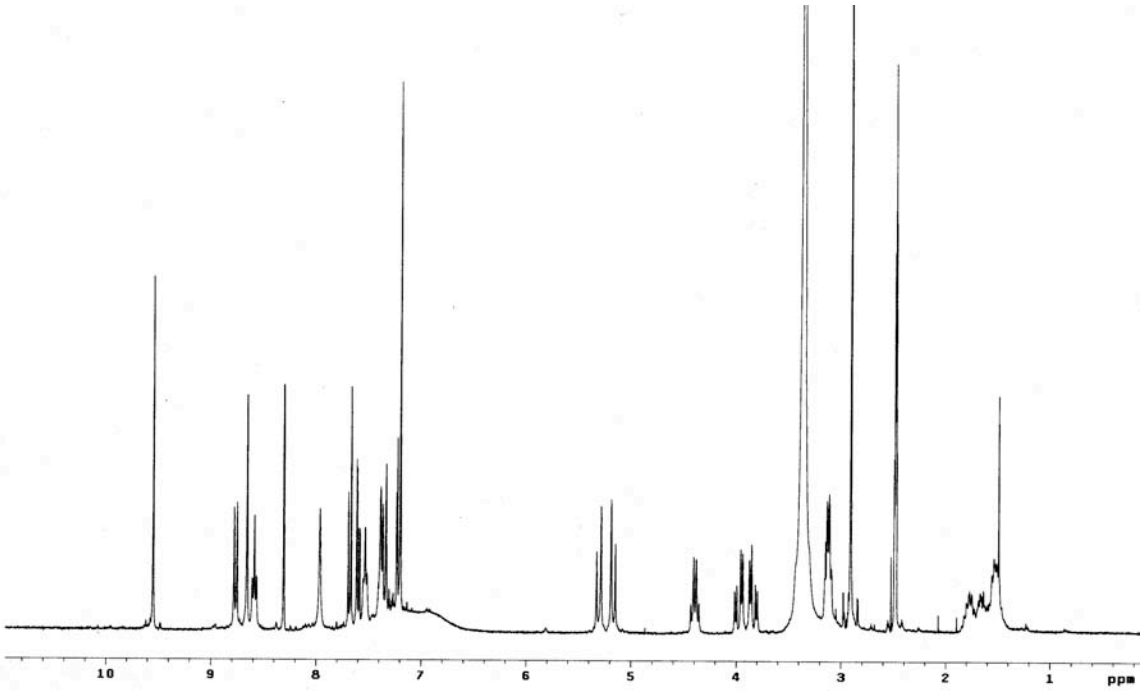
^{13}C NMR (75 MHz, DMSO-d6)

 = 172.4, 168.6, 167.9, 166.8, 157.2, 151.9, 137.9, 135.8, 132.0, 131.2, 130.1, 127.6, 126.0, 124.5, 123.6, 123.2, 121.9, 115.2, 68.8, 54.1, 43.9, 27.6, 26.2

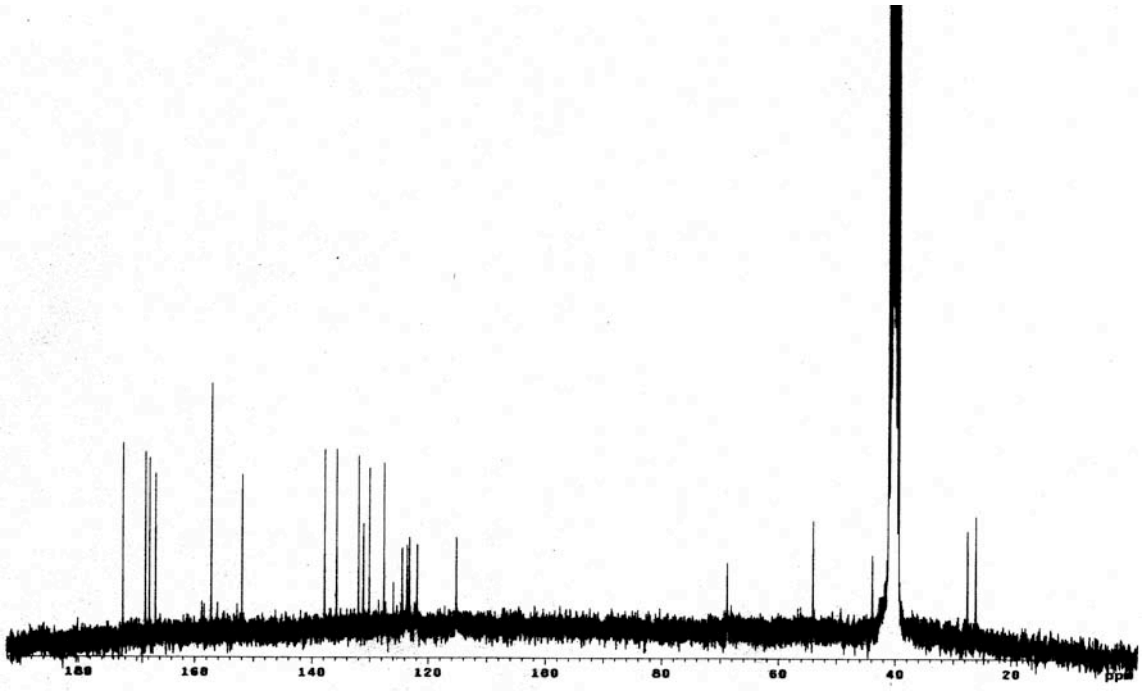
MS (MALDI) calc'd for $C_{24}H_{30}N_8O_7S$ ($M+H^+$) 575.6, found 575.1

Analytical HPLC

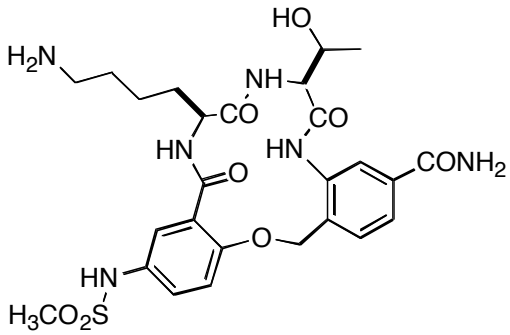
homogeneous single peak, retention time = 18 min (2-40% B in 30 min)



^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)



37e

^{1}H NMR (300 MHz, DMSO-d_6)

δ = 9.65 (s, 1H), 8.85 (d, *J* = 8.4 Hz, 1H), 8.73 (s, 1H), 8.36 (d, *J* = 1.91 Hz, 1H), 8.01 (s, 1H), 7.88 (d, *J* = 8.4 Hz, 1H), 7.70-7.51 (m, 4H), 7.43 (m, 2H), 7.32 (m, 2H), 7.26 (d,

J = 3.0 Hz, 1H), 5.47 (d, *J* = 12.3 Hz, 1H), 5.03 (d, *J* = 12.3 Hz, 1H), 4.37-4.22 (m,3H), 2.97 (s, 3H), 2.81-2.70 (m, 2H), 1.80 (m, 2H), 1.59 (m, 2H), 1.45 (m, 2H), 1.10 (d, *J* = 6.6, 3H)

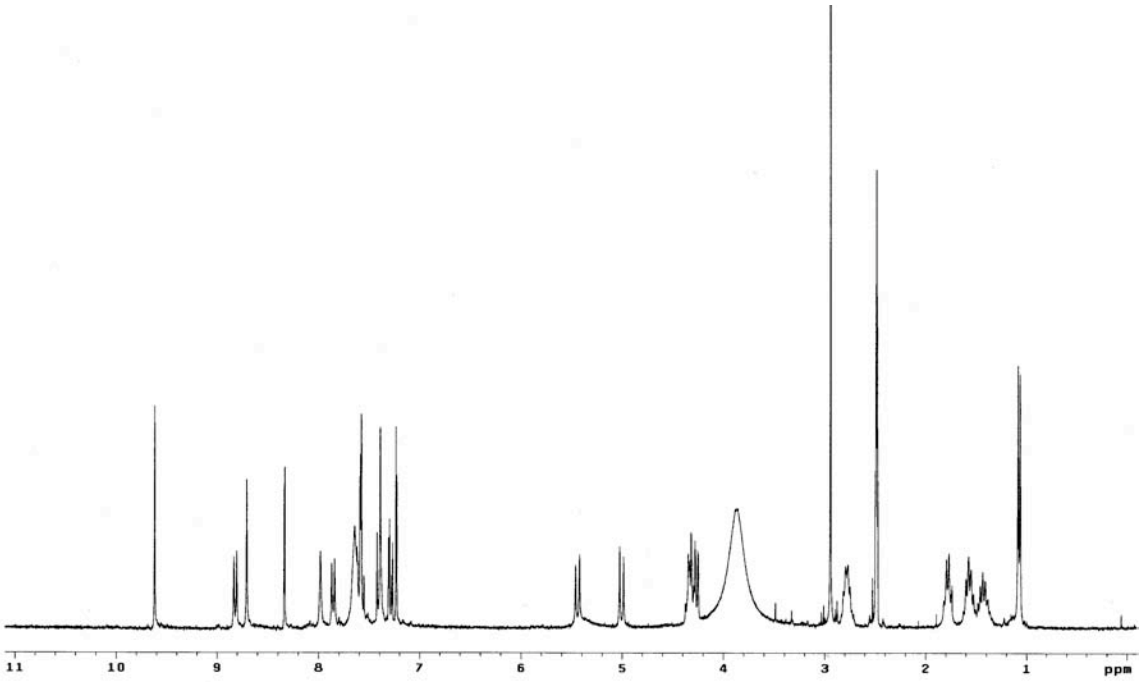
^{13}C NMR (75 MHz, DMSO-d_6)

 = 173.7, 169.9, 167.9, 166.8, 151.7, 137.8, 135.9, 132.2, 130.8, 130.0, 127.5, 124.6, 123.8, 123.2, 122.0, 114.7, 69.0, 66.3, 59.8, 56.0, 30.8, 27.2, 23.5, 21.2

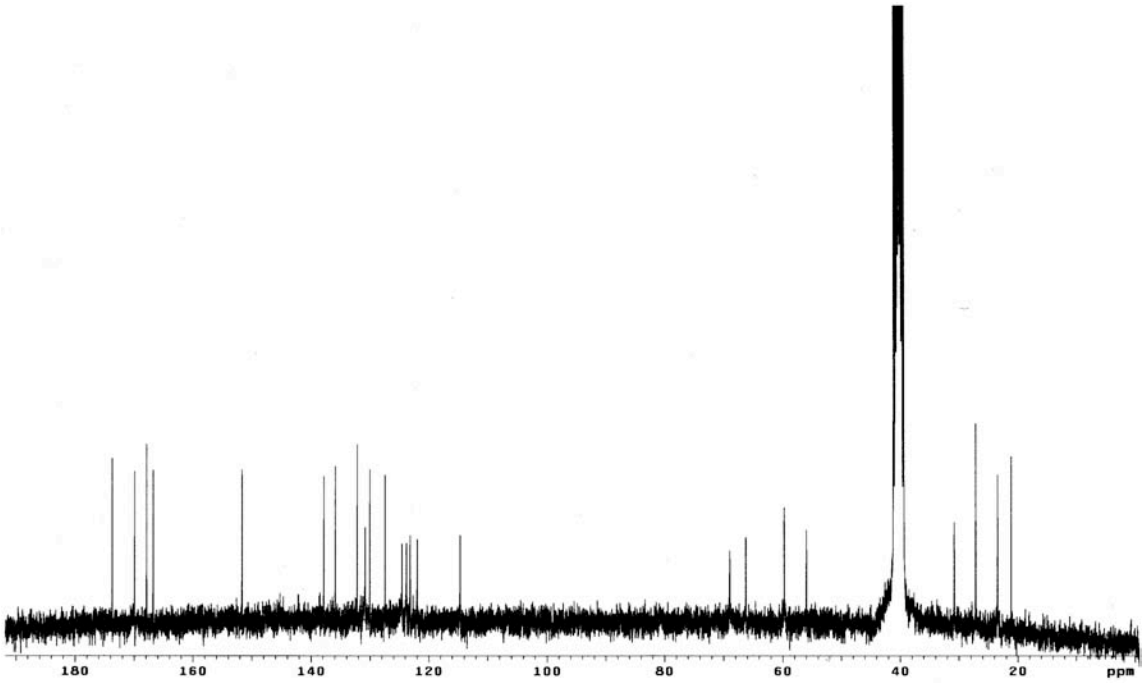
LRMS (MALDI) calc'd for $C_{26}H_{34}N_6O_8S$ ($M+H^+$) 591.6, found 591.1

Analytical HPLC

homogeneous single peak, retention time = 12.7 min (2-40% B in 30 min)

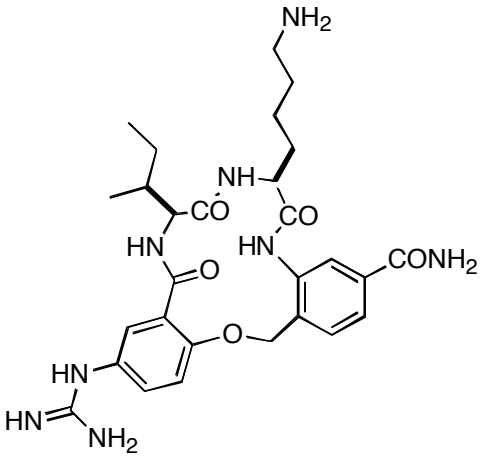


^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

Guanidine Derivatives 38a-c. Resin containing side-chain protected cyclic peptide with amino group was swelled in CH_2Cl_2 in a fritted syringe for 30 min and treated with 10 equiv. of *N, N'-bis*-BOC-1-guanylpyrazole in DMF. After gental shaking for 24 h, the resin was washed with DMF (3x), MeOH (3x), CH_2Cl_2 (3x), DMF (2x), MeOH (2x), and CH_2Cl_2 (3x) and the peptide was cleaved from the resin by treatment with a mixture of 90% TFA, 5% H_2O and 5% TIS for 2 h. The cleavage solution was separated from the resin via filtration. After most of the cleavage cocktail (about 90%) was evaporated in a stream of nitrogen, the crude peptide was triturated using anhydrous ethyl ether. The purity of this crude product was determined by analytical HPLC (SSI system, 2-40% B in 30 min). The crude product was dissolved in H_2O, filtered, then purified with preparative HPLC and finally lyophilized to yield products **38a-c**.



38a

^{1}H NMR (500 MHz, DMSO-d6)

δ = 9.64 (s, 1H), 8.78 (d, *J* = 8.4 Hz, 1H), 8.76 (s, 1H), 8.52 (d, *J* = 9.3 Hz, 1H), 8.40 (s, 1H), 7.97 (s, 1H), 7.75-7.58 (m, 4H), 7.49 (d, *J* = 8.7 Hz, 1H), 7.38 (m, 5H), 7.30 (m, 1H), 5.51 (d, *J* = 12.5 Hz, 1H), 5.27 (d, *J* = 12.5 Hz, 1H), 4.40 (m, 1H), 4.28 (t, *J* = 9.3 Hz, 1H), 2.82-2.70 (m, 2H), 1.95 (m, 2H), 1.76 (m, 1H), 1.57 (m, 3H), 1.40 (m, 2H), 1.21 (m, 1H), 0.98-0.82 (m, 6H)

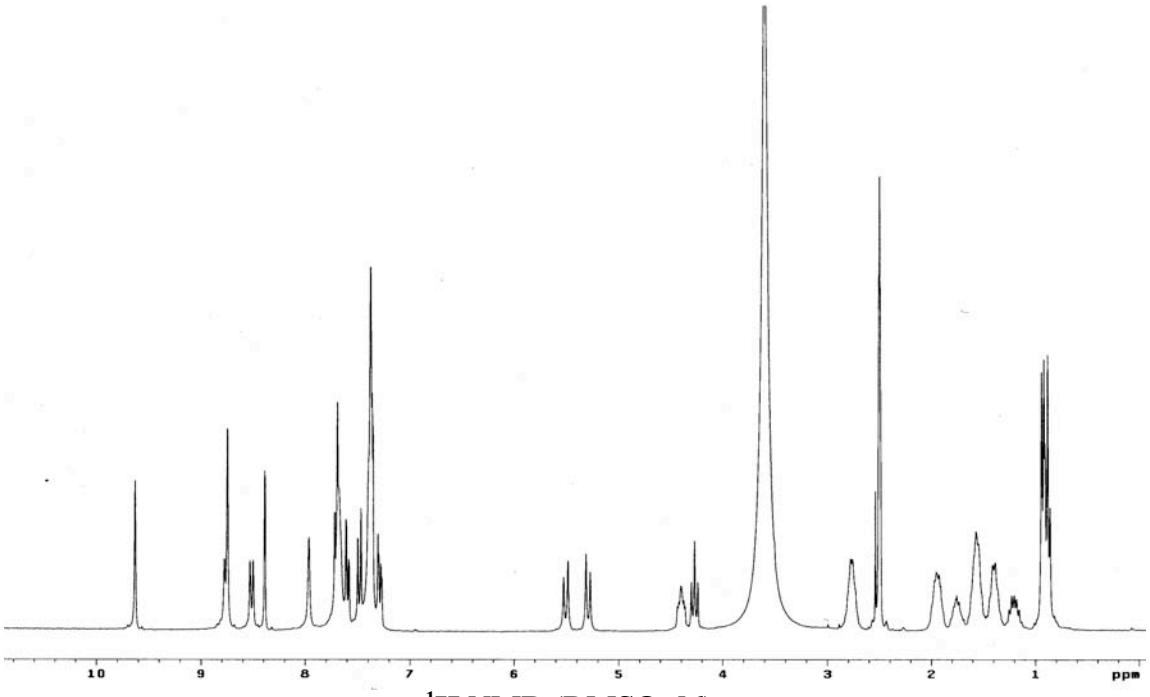
^{13}C NMR (75 MHz, DMSO-d6)

δ = 172.3, 171.0, 167.9, 165.7, 156.8, 153.4, 137.8, 136.1, 131.4, 129.6, 129.3, 128.7, 127.5, 126.6, 123.5, 122.8, 115.0, 67.9, 60.7, 53.7, 35.7, 30.9, 27.1, 26.0, 23.3, 16.3, 11.4

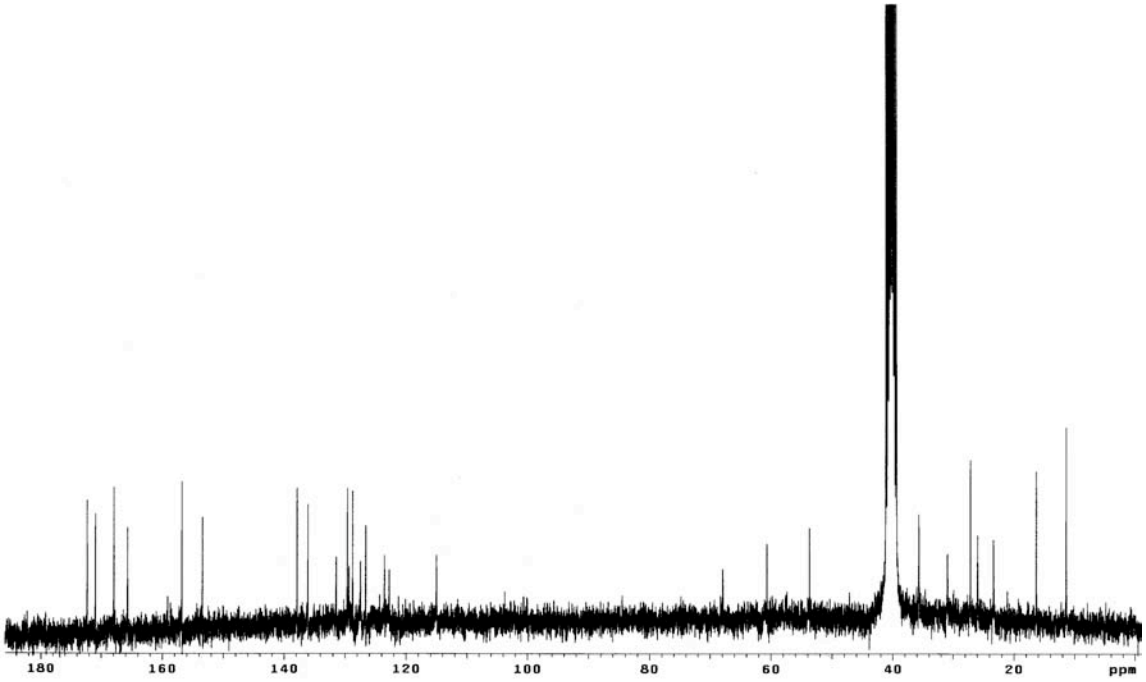
LRMS (MALDI) calc'd for $C_{28}H_{38}N_8O_5$ ($M+H^+$) 567.6, found 567.1

Analytical HPLC

homogeneous single peak, retention time = 13.7 min (2-40% B in 30 min)



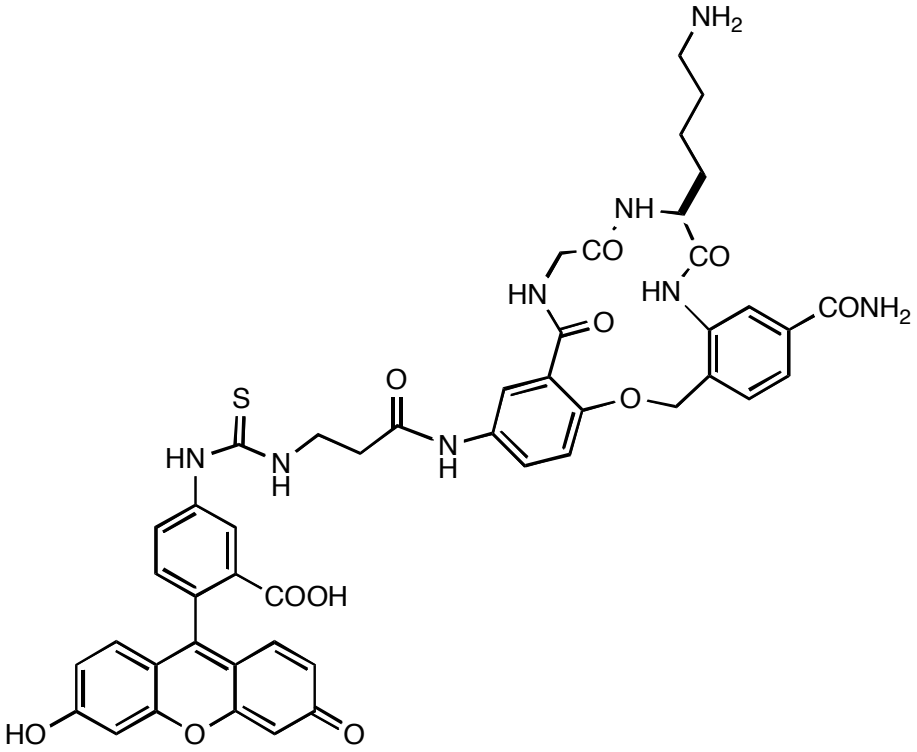
^{1}H NMR (DMSO-d6)



^{13}C NMR (DMSO-d6)

General Procedure for Preparation of the Fluoresceinylated Macrocycles 41

Resin containing side-chain protected cyclic peptide with amino group was swelled in CH_2Cl_2 in a fritted syringe for 30 min and treated with 4 equiv Fmoc-β-Ala-OH, 4.6 equiv PyBrOP and 15 equiv 2,6-lutidine in CH_2Cl_2. After 20 h shaking, the resin was washed and the Fmoc protecting group was removed with 20% piperidine in DMF. Fluorescein 5-isothiocyanate, (*ie* "isomer I"; 3 equiv) was then added with 5 equiv DIEA in CH_2Cl_2/DMSO (2:1) for 3 h. After washing, the resin was cleaved with a mixture of 90% TFA, 5% TIS, and 5% H_2O for 2 h. After most of the cleavage cocktail was evaporated in a stream of nitrogen. Anhydrous ethyl ether was added to precipitate the product. The crude product was purified by preparative HPLC and lyophilized to yield final products **41a-e**.



41b

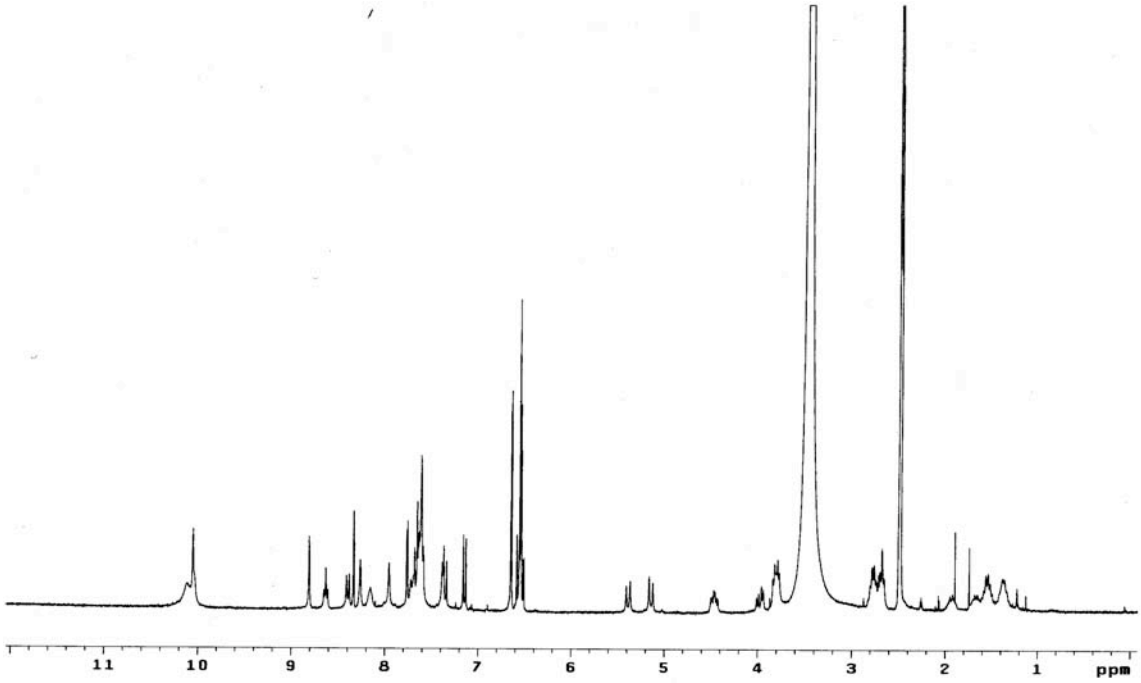
^{1}H NMR (500 MHz, DMSO-d6)

 = 10.14 (bs, 1H), 10.08 (s, 1H), 8.83 (s, 1H), 8.66 (t, *J* = 6 Hz, 1H), 8.42 (d, *J* = 8.7 Hz, 1H), 8.36 (s, 1H), 8.29 (s, 1H), 8.18 (bs, 1H), 7.98 (s, 1H), 7.79 (d, *J* = 2.64 Hz, 1H), 7.77 – 7.59 (m, 6H), 7.42-7.36 (m, 2H), 7.17 (d, J = 8.7 Hz, 1H), 6.67 (d, *J* = 2,19 Hz, 3H), 6.62 – 6.52 (m, 5H), 5.41 (d, J = 12.3 Hz, 1H), 5.16 (d, J = 12.3 Hz, 1H), 4.48 (m, 1H), 4.40 (m, 2H), 3.83 (m, 2H), 2.79 (m, 2H), 2.70 (m, 2H), 1.94 (m, 1H), 1.69 (m, 1H), 1.57 (m, 2H), 1.39 (m, 2H)

LRMS (MALDI) calc'd for $C_{47}H_{44}N_8O_{11}S$ ($M+H^+$) 929.9, found 929.2

Analytical HPLC

98.2 %, retention time = 15.8 min (10-50% B in 30 min)



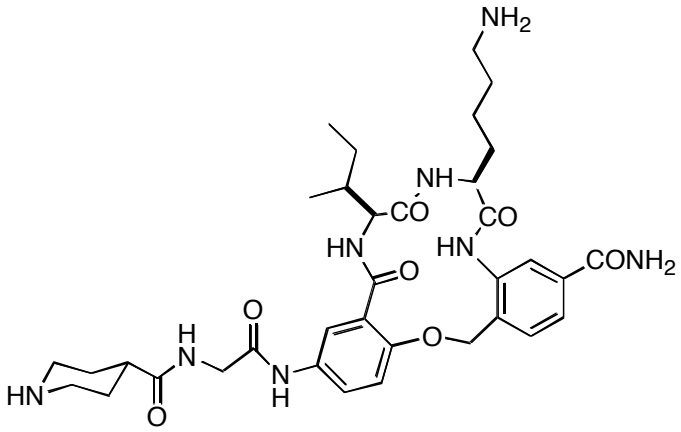
^{1}H NMR (DMSO-d6)

General Procedure for Preparation of Monomeric Precusors 42 and 43

Resin containing side-chain protected cyclic peptide with amino group was swelled in CH_2Cl_2 in a fritted syringe for 30 min and treated with 4 equiv Fmoc-Gly-OH, 4.6 equiv PyBrOP and 15 equiv 2,6-lutidine in CH_2Cl_2. After 30 h shaking, the resin was washed and the Fmoc protecting group was removed with 20% piperidine in DMF. The resin was then washed and treated with 3 equiv Fmoc-Inp-OH, 5equiv DIC and 5 equiv HOBt in DMF for 8 h. The resin was washed and Fmoc was removed. The resin was washed again, dried *in vacuo* and split into two portions. One portion was subjected to cleavage condition via the treatment with a mixture of 90% TFA, 5% TIS and 5% H_2O for 2 h. After most of the cleavage cocktail was evaporated in a stream of nitrogen. Anhydrous ethyl ether was added to precipitate the product. The crude product was purified by preparative HPLC and lyophilized to yield final products **42a-c**.

The other portion of the resin was swelled in CH_2Cl_2 in a fritted syringe for 30 min and treated with 2 equiv dichlorotriazinylaminofluorescein, 3 equiv $^iPr_2NEt$ in a

mixture of DMSO/$CHCl_3$ (1:3) for 3 h. The resin was washed and subjected to the above cleavage condition. After most of the cleavage cocktail was evaporated in a stream of nitrogen. Anhydrous ethyl ether was added to precipitate the product. The crude product was purified by preparative HPLC and lyophilized to yield final fluorescently-labeled products **43a-c**.



42a

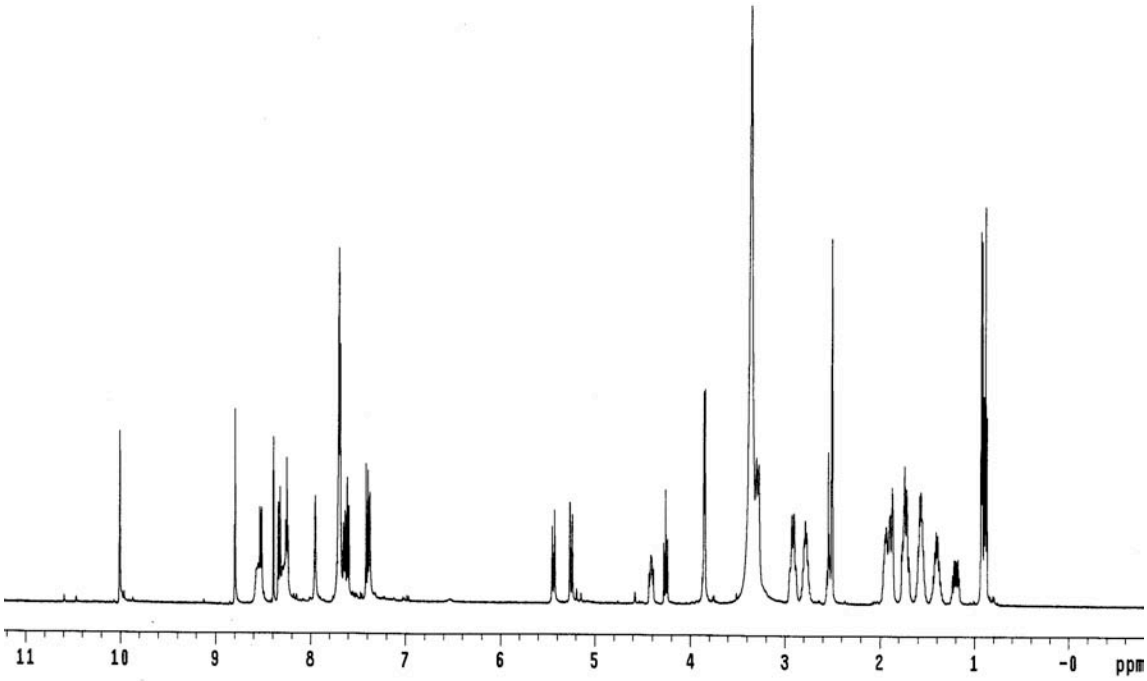
^{1}H NMR (500 MHz, DMSO-d6)

 10.01(s, 1H), 8.8 (s, 1H), 8.53 (d, J = 8.0, 1H), 8.39 (s, 1H), 8.33 (d, J=9.0), 8.29 (m, 1H), 8.25 (t, J=5.5, 1H), 7.95 (s, 1H), 7.73-7.67 (m, 4H), 7.64 (dd, J=9.0, 2.6, 2H), 7.61(dd, J=8.5, 1.5, 2H), 7.42-7.36 (m, 2H), 5.44 (d, J = 12.0, 1H), 5.26 (d, J =12.0, 1H), 4.45-4.38 (m, 1H), 4.26 (t, J=9.0, 1H), 3.85 (d, J=6.0, 2H), 3.33-3.25 (m, 2H), 2.98-2.87 (m, 2H), 2.84-2.73 (m, 2H), 2.57-2.51 (m, 1H), 1.99-1.84(m, 4H), 1.78-1.68 (m, 3H), 1.62-1.52 (m, 3H), 1.46-1.34 (m, 2H), 1.25-1,15 (m, 1H), 0.95-0.86 (m, 6H)

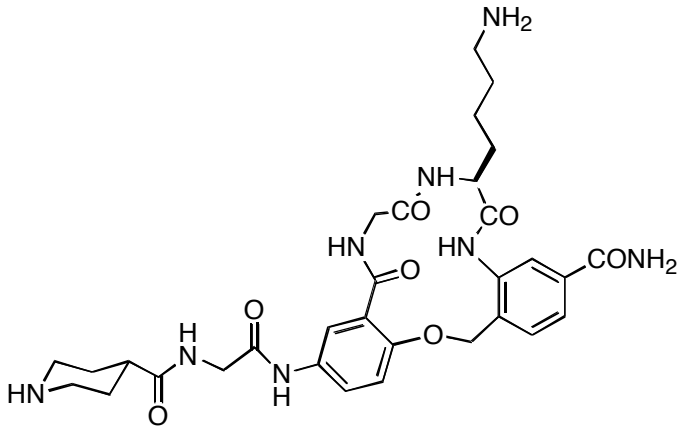
MS (MALDI) calc'd for $C_{35}H_{48}N_8O_7$ ($M+H^+$) 693.8, found 693.2

Analytical HPLC

homogeneous single peak, retention time = 14.3 min (2-40% B in 30 min)



^{1}H NMR (DMSO-d6)



42b

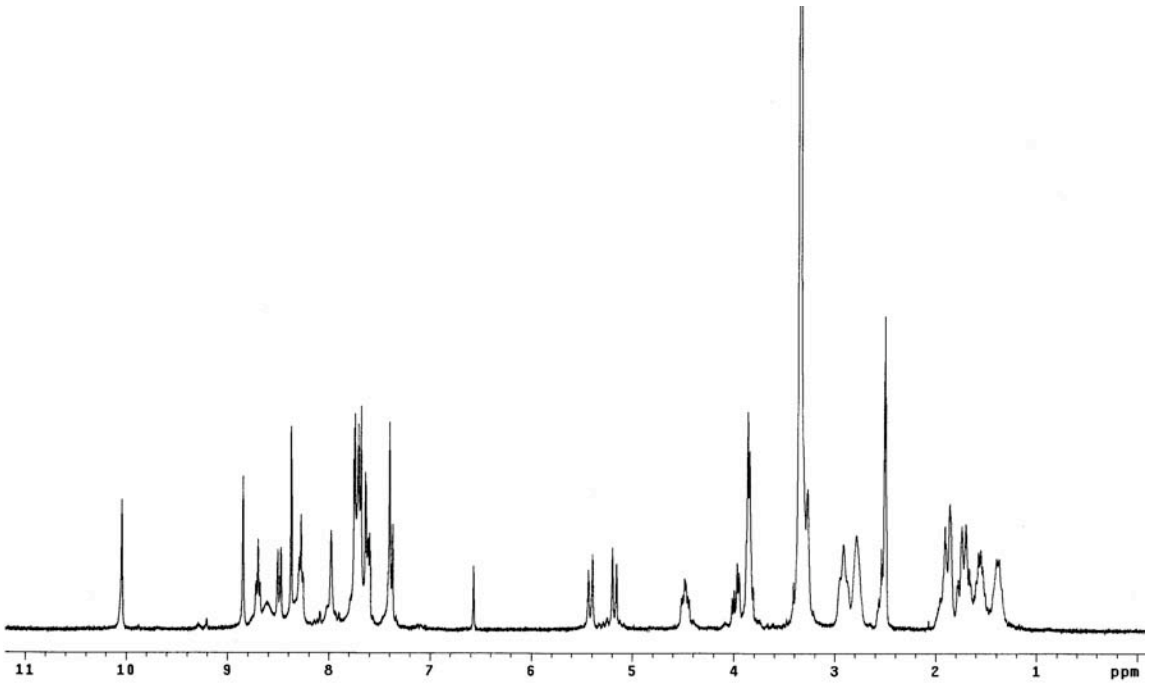
^{1}H NMR (500 MHz, DMSO-d6)

 10.05(s, 1H), 8.87 (s, 1H), 8.72 (t, J = 5.5, 1H), 8.50 (d, J=8.5, 1H), 8.40 (s, 1H), 8.28 (t, J=4.65, 1H), 7.99 (s, 1H), 7.80-7.61 (m, 7H), 7.44-7.38 (m, 2H), 5.43 (d, J = 12.0, 1H), 5.20 (d, J =12.0, 1H), 4.54-4.47 (m, 1H), 4.01(dd, J=6.5, 6.5, 1H), 3.91-3.83 (m, 3H), 3.33-3.27 (m, 2H), 2.94 (t, J=9.5, 2H), 2.85-2.76 (m, 2H), 2.59-2.52 (m, 1H), 2.01-1.85 (m, 3H), 1.81-1.66 (m, 3H), 1.65-1.51 (m, 3H), 1.48-1.34 (m, 2H)

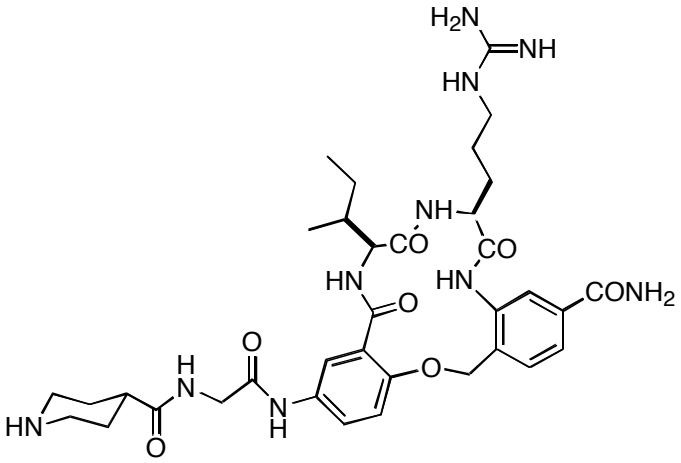
MS (MALDI) calc'd for $C_{31}H_{40}N_8O_7$ ($M+H^+$) 637.7, found 637.2

Analytical HPLC

homogeneous single peak, retention time = 11.3 min (2-40% B in 30 min)



^{1}H NMR (DMSO-d6)



42c

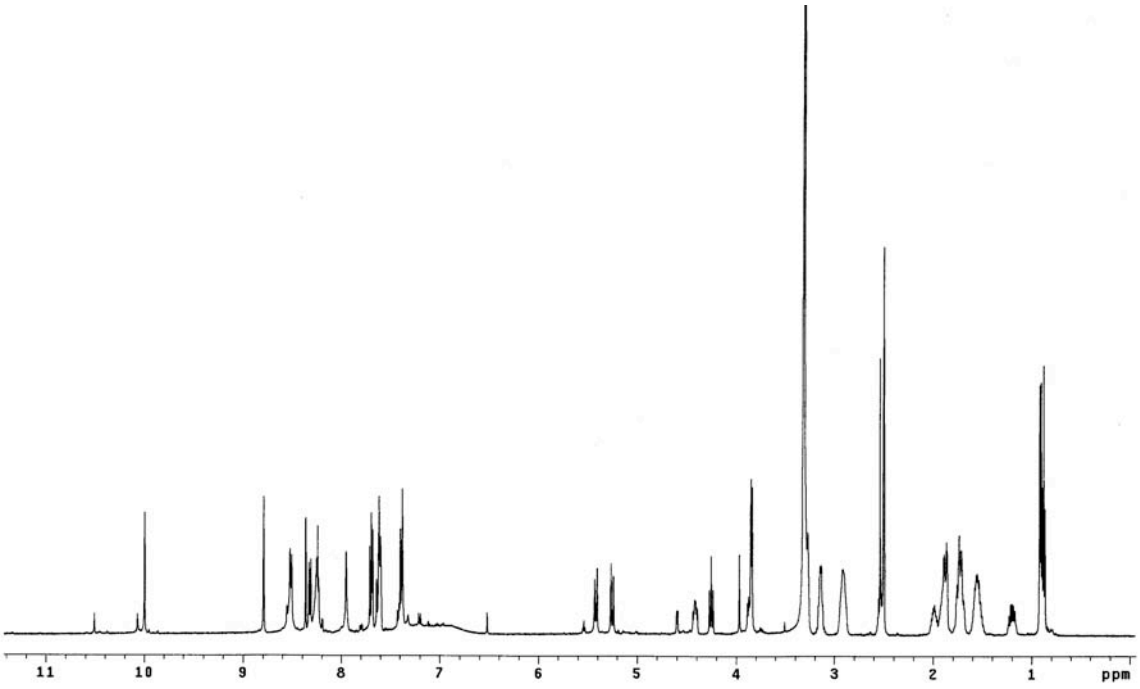
^{1}H NMR (500 MHz, DMSO-d6)

 10 (s, 1H), 8.79 (s, 1H), 8.57-8.49 (m, 2H), 8.37 (s, 1H), 8.33 (d, J=9.0, 1H), 8.29-8.22 (m, 2H), 7.96 (s, 1H), 7.72-7.67 (m, 2H), 7.65-7.59 (m, 3H), 7.42-7.36 (m, 3H), 5.42 (d, J = 12.0, 1H), 5.25 (d, J =12.0, 1H), 4.45-4.38 (m, 1H), 4.26 (t, J=9.0, 1H), 3.85 (d, J=6.0, 2H), 3.31-3.26 (m, 2H), 3.18-3.11 (m, 2H), 2.96-2.86 (m, 2H), 2.57-2.51 (m, 1H), 2.05-1.95(m, 1H), 1.95-1.84 (m, 3H), 1.79-1.66 (m, 3H), 1.62-1.48 (m, 3H), 1.25-1,15 (m, 1H), 0.94-0.85 (m, 6H)

MS (MALDI) calc'd for $C_{35}H_{48}N_{10}O_7$ ($M+H^+$) 721.8, found 721.3

Analytical HPLC

93.6 %, retention time = 15.4 min (2-40% B in 30 min)



^{1}H NMR (DMSO-d6)

VITA

Name:		Zebin Xia
Permanent Address:		Yan Shan Cun 4-502 Hunan University Changsha 410082, P.R.China
Education:	1998-2004	Ph.D., ITDE, Texas A&M University College Station, TX
	1988-1991	M.S., Organic Chemistry Nankai University Tianjin, P.R.China
	1982-1986	B.S., Chemistry, Hunan Normal University Changsha, P.R.China